

How not to be charitable

A scandal in France reveals that large sums of money intended for cancer research have been misdirected. Scientists and government alike have failed colleagues and the public.

FEW experiences are more wounding than to see a generous and selfless gift selfishly misused by its recipient. So the revelations last week that a major French cancer charity has been bestowing large sums in highly questionable ways (see page 103) will hurt many people: the donors of the funds, innocent recipients of research funds who may be contaminated by association, applicants for funds who were wrongly rejected, and, not least cancer sufferers.

The allegations of serious mismanagement at France's biggest medical charity, L'Association pour la Recherche sur le Cancer (ARC), made by the national audit commission, are serious. But what is even more shocking is that it has taken more than a decade for the situation to be officially recognized. Rumours about the running of the ARC have circulated for at least 15 years. But until last week the organization's donors, who annually contribute around FF600 million, have had no clear idea of how their money was spent.

The answer, it now appears, is "badly". According to the audit, only a quarter of the charity's spending has gone directly on research. To make things worse, many of the grants distributed by the ARC were not assessed through proper peer review.

The facts did not come to light sooner because the ARC has fiercely resisted outside inspection. The organization has, on occasion, taken advantage of researchers and physicians who benefited from its largesse, in order to challenge accusations of mismanagement. When allegations similar to those of the audit commission were made in 1994, the ARC solicited letters of support from researchers — more than a thousand of whom obliged. The letters were fed into the ARC's public relations machine, which had been put into high gear in a desperate bid to alleviate public concern.

Proper management of such charities is doubly important because of their growing importance in research. The national biomedical research organization INSERM receives about FF75 million annually from charities. While this amount is small compared with INSERM's total budget of FF2.7 billion, it is significant compared with the FF300 million provided by INSERM to its laboratories for equipment and supplies. Some INSERM laboratories depend on the ARC for up to half of such spending.

Arguably, some of the ARC's current woes stem from its origins. It was created by one man, Jacques Crozemarie, an engineer at the Centre National de la Recherche Scientifique (CNRS), following the death of his wife from a brain tumour. The charismatic Crozemarie has a flair for sophisticated fund-raising techniques which resulted in the collection of large sums, which in turn gained Crozemarie the support of many prominent cancer researchers. The agency has since evolved as a mammoth self-appointed fund-raising agency, to which donors gave in the belief that their money would find its way to cancer research in one way or another.

This mutual dependency between Crozemarie and leading cancer researchers appears to have resulted in an autocratic organization of the charity that has given a few individuals control of the distribution of enormous sums of money. The executive board includes many members who receive funds from the charity. Com-

bine secrecy and conflict of interest in that way, and you maximize the potential for abuse. The charity must be reconstituted.

The scandal highlights a glaring lack of regulation of charities in France — there is no organization capable of insisting on adequate standards of openness and scientific integrity. Policy-makers in other countries need to take note of this debacle and reflect on whether their regulatory environment is equally lax. If so, prompt audits of research charities may be advisable.

Researchers involved in the running of charities and non-profit organizations everywhere should also be scrutinizing their consciences. Even if their formal accountabilities are private, there is a clear public duty that they should achieve, and be seen to achieve, exemplary levels of transparency and quality. The example provided by scientists on the board of the ARC appears to be much less than inspiring. □

Money is not enough

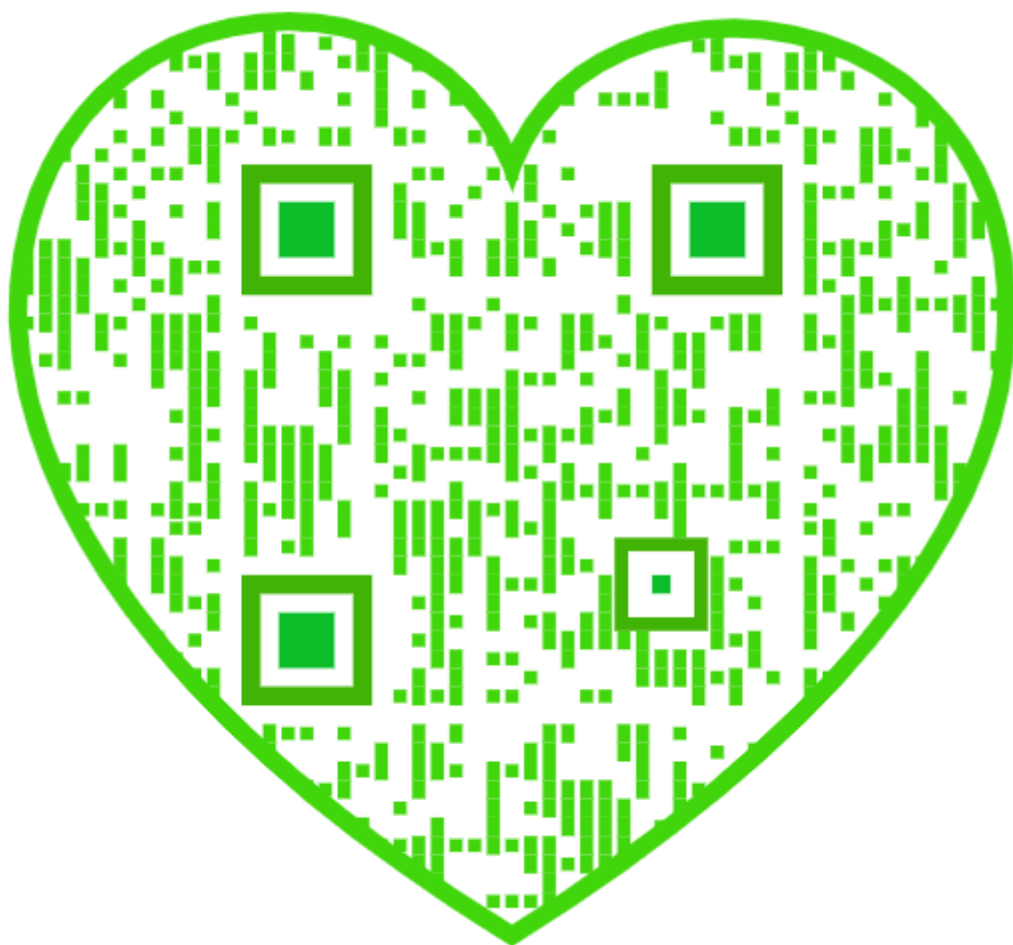
The increases in Japan's budgets for science are welcome. But better management and supervision are also required.

THE Japanese government at last seems to be making efforts to pump much needed extra money into its public-sector research system. The increases in science-related budgets (see page 105) are remarkable in Japan's troubled economic circumstances. Much of the credit can be given to a small band of comparatively young politicians who are lobbying for science and have passed a new law to help them to that end (see *Nature* 378, 227; 1995). But they should also ensure that the money is well spent.

The complex and bureaucratic dispersion of public research funds is extremely wasteful of Japanese taxpayers' money. For example, universities often have plenty of funds to buy equipment but have no money to employ technicians to look after it. Gleaming new machines have to be maintained by inexperienced graduate students (who deserve more creative tasks), while older — but still functional — equipment is left to rot. More money has to be spent on technicians and on offering them salaries that are competitive with industry.

Also lacking is a system of supervision to ensure that government projects represent sensible investments. The past few decades are strewn with examples of wasteful government projects that led virtually nowhere — the nuclear-powered ship *Mutsu*, for example, gobbled up more than a billion dollars and spent all but a few days of its 15-year life sitting in port. Now it is being converted into a diesel-powered oceanographic vessel at considerable cost to Japan's taxpayers. The notorious fifth-generation computer and the Monju fast-breeder reactor are other underwhelming examples of the country's use of public funds.

Fortunately, the same band of young politicians that is lobbying for more money is also trying to set up a science and technology assessment organization to oversee government projects (see *Nature* 378, 657; 1995). It is to be hoped that it succeeds and that the organization is truly impartial and well-informed in its judgments. □



Agencies warn of grant delays as Washington returns to backlog

Washington. A three-week shutdown of large parts of the US federal government ended this week, leaving science agencies warning of months of delays and uncertainty as their administrative staff returned to face huge backlogs of grant processing and other work.

Nature compounded the shambles at the Washington headquarters of the National Institutes of Health (NIH), the National Science Foundation (NSF) and the National Aeronautics and Space Administration (NASA) on Monday, when 18 inches of snow pushed the anticipated reopening back to the middle of this week.

The unprecedented three-week shutdown ended with a retreat by Republican congressmen, who were being blamed by the public for the disruption that it was causing. Last Friday, 5 January, the Republicans

changed their strategy and passed bills to reopen the government, despite failing to reach agreement with President Bill Clinton on how to balance the government's budget.

For the NIH, the shutdown ended with an unexpected financial windfall. One of three 'continuing resolution' bills passed into law at the weekend will assure the biomedical research agency of funding to the end of the fiscal year, 30 September, of almost \$11.5 billion, up by 5.7 per cent from last year.

NIH and the Centers for Disease Control, based in Atlanta, Georgia, are included in a bunch of 'targeted programmes' which Congress decided should no longer be held hostage to budget negotiations. NSF and NASA failed to make the list, and are assured of funding — at the levels of the last financial year — only until 25 January, when the threat of yet another shutdown looms.

Grant processing at all of these agencies has been at a standstill now for almost a month, holding up hundreds of millions of dollars of funds for university research. At the NIH, 2,000 grants due to start or be renewed on 1 December and 1 January have not been issued. NIH officials say they are uncertain how quickly the return to work will enable the processing of these and a further 2,000 grants due on 1 February.

Harold Varmus, the director of the NIH, says even with assured funding for the rest of the year, the shutdown will have "a very significant fallout". NIH has "a very tightly orchestrated system" to process 30,000 grants a year. He appealed to grant recipients to be patient with NIH staff as they seek to get the process back on schedule.

Neal Lane, director of the NSF, says that the agency, which funds most non-biomedical university science in the United States, has delayed \$100-million worth of awards, and that this has been growing at a rate of \$10 million a day. "There's no way that we'll get through the backlog quickly," says Lane, who noted last Friday there were 42 large carts of mail in his mailroom.

After the agency failed to make the list of targeted programmes, another senior NSF official attacked universities for failing to complain loudly enough on its behalf. The NSF got "zero help" from academics in its bid to make the list, the official said. Asked how quickly NSF would clear its backlog, the official pointed to the looming threat of another shutdown on 25 January. "I don't know what we can do 'quickly,'" he said. "We've got a two-week window to do ten weeks of work."

Most major universities have tried to ease the impact of the shutdown on researchers by providing temporary funds to enable the continuation of research for which grant support has been delayed. The Massachusetts Institute of Technology, for example, had last week committed \$11.5 million a year, or about \$1 million a month, to supporting such projects.

Mike Amey, assistant dean for research administration at the school of medicine at Johns Hopkins University, Baltimore, says the impact of the three-week shutdown is hard to calculate, because of the holiday season and the normally slow start to the financial year. "It's not that we are not feeling the pinch," he says. "But it is too soon to know how big the impact is, as there is normally a delay at this time of year."

NASA employees also returned to a huge backlog of paperwork this week, which ►

CERN greets detection of first anti-atom

London. **An announcement last week by the European Laboratory for Particle Physics (CERN) that its physicists have created anti-hydrogen — the first detected anti-atom — has set the stage for a new type of experiment designed to test fundamental physical principles.**

Those tests, which still require further experimental development, will cause an even greater stir if, contrary to most physicists' expectations, the principles — namely the equivalence principle (EP) and universal compliance with charge-parity-time (CPT) symmetry — are overturned.

The CERN physicists, headed by Walter Oelert of the Institute of Nuclear Physics at the National Research Centre in Jülich, brought together antiprotons and positrons in CERN's Low Energy Antiproton Ring (LEAR, above right). By studying the resulting particles ejected from the ring, the researchers were able to confirm that anti-hydrogen had been formed, albeit fleetingly.

The next challenge, to be pursued in other facilities, will be to carry out extremely precise spectroscopic measurements of the energy levels of anti-atoms, in order to look for differences between hydrogen and

anti-hydrogen spectra.

Such spectroscopy requires the development of anti-atom 'traps'. This is unlikely to be achieved at CERN before late 1998 — and then only if physicists can find about \$Fr7 million (US\$6 million) from external sources required to develop the necessary equipment. Meanwhile, physicists at Fermilab are hoping to carry out even more challenging spectroscopy of beams of anti-hydrogen atoms within a year or so.

Philip Campbell

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

David Parker/SPL

► will cause delays in processing some 2,000 research grants supporting university and federal scientists. Researchers whose NASA grants are scheduled for renewal in the first quarter of 1996 can expect delays in receiving their money, according to Henry Brinton of the agency's space science office.

Some spacecraft launches may also slip. Work on the TIMED (Thermosphere, Ionosphere, Mesosphere Energetics and Dynamics) spacecraft is already three months late in getting started, as start-up funding is contained in the stalled 1996 appropriations bill.

That makes a slip in the launch date likely, says Stamatios Krimigis of the Johns Hopkins University Applied Physics Laboratory (APL), which is building the spacecraft. Other 'fast, cheap' spacecraft missions with tight deadlines may also have trouble meeting their schedules.

Two science missions scheduled for launch in February are expected to remain on schedule, however. APL's Near Earth Asteroid Rendezvous (NEAR), the first of NASA's small Discovery planetary missions, has adequate funding up to its launch, says Krimigis, as does the POLAR spacecraft that will study the Earth's magnetosphere.

NASA archival data have remained available to researchers through the Internet during the shutdown, which did not affect contractors working at the National Space Science Data Center at NASA's Goddard Space Flight Center in Maryland. NASA employees have also been allowed to work on key projects such as the Galileo mission to Jupiter and the X-Ray Timing Explorer spacecraft, launched on 30 December.

But Galileo scientists have been frustrated at being unable to release results from the probe's entry into Jupiter's atmosphere on 7 December. A press conference planned for 19 December had to be cancelled as several of the principal investigators for the probe are NASA civil servants.

Employees at the Jet Propulsion Laboratory in Pasadena, California, were spared most of the effects of the furlough, as they are employed by the California Institute of Technology, rather than directly by NASA. A threat of extensive lay-offs there was lifted with last weekend's deal in Washington to reopen the government.

Colin Macilwain & Tony Reichhardt

Systems analysis institute names new director

Vienna. Gordon Macdonald, currently professor of international relations at the University of California in San Diego, and formerly chief scientist of the MITRE Corporation in Bedford, Massachusetts, has been appointed the next director of the International Institute for Applied Systems Analysis, based at Laxenberg, near Vienna.

Macdonald, who will begin his three-year term of office in August, has chaired numerous scientific committees for the White House and various federal agencies,

AT&T rings out redundancies for Bell Labs' researchers

Washington. AT&T Bell Laboratories — the largest and most famous industrial research laboratories in the United States whose activities have already been under pressure for several years — are facing yet further retrenchment as the result of an announcement by AT&T last week that it is to cut 40,000 of its 300,000 employees.

Scientists at the laboratories' main site at Murray Hill, New Jersey, say that morale among the 6,000 staff there, already low as a result of earlier redundancies, has fallen still further. A round of voluntary redundancies at the end of last year failed to attract enough volunteers, leading to an unspecified number of involuntary sackings.

"It's like living through a bombing raid," said one physicist, who declined to be identified. "They are demolishing the best place ever [for doing research]." The recent redundancies will go some way to meeting the laboratories' share of the 40,000 job cuts. But a spokesman for the laboratories said that more would be needed in light of the latest announcement.

The AT&T Bell Laboratories network employs about 25,000 people, including 4,000 PhDs, around the world. Most of the 25,000 are engaged in product development for AT&T. But the laboratories still spend around \$500 million a year on a basic research programme, which is based entirely in the United States, and supports 2,000 principal investigators and their staff.

Even before last week's announcement, the laboratory network, which includes sites in eight US states and 21 smaller outposts abroad, was in the process of being divided into two components as part of a plan announced last September to split up the giant AT&T corporation.

Sam Bleecker, a spokesman for AT&T Bell Laboratories, said that the corporation was not ready to comment on the impact that the new announcement would have on the laboratories, which he said was expected to be "minimal". But he admitted that "there will be some lay-offs, not just

on the technical side but in all areas".

Bleecker said that it would be "premature" for management of the laboratories to talk about their future at this stage, adding that an announcement would be made "when the situation is less fluid".

The AT&T corporation, which has an annual turnover of \$75 billion, will split later this year into three parts. One will be a long-distance telephone corporation, which will keep the AT&T name and will have a turnover of \$50 billion. The second will be a systems and technology corporation, as yet unnamed, which will manufacture telephone exchanges and other high technology equipment, with an annual turnover of \$20 billion. And the third component will be a relatively small computer corporation.

The larger part of AT&T Bell Laboratories will re-adopt the name of Bell Laboratories, and will be part of the systems and technology corporation which, in a move intended to symbolize its commitment to technical innovation, will establish its headquarters at Murray Hill. But a substantial part of the laboratory network will be transferred to a new organization, to be called AT&T Laboratories, inside the telephone corporation.

According to sources at Murray Hill, AT&T intends to transfer about two-fifths of its \$500-million basic research programme — including teams working in systems design, software and applied mathematics — into the telephone corporation. That will leave \$300 million of basic research at Bell Laboratories, supporting a systems and technology corporation with annual sales of \$20 billion.

Apart from the immediate threat of redundancies, scientists fear that this amount of basic research will be too much for the new corporation to carry. "Being a tax on a \$75 billion corporation is better than being a tax on a \$20 billion corporation," says one. AT&T says that the systems and technology company will need a strong research programme. But a \$300-million programme, at 1.5 per cent of turnover, would be substantially above the industry norm of 1 per cent.

Bell Laboratories' emphasis has shifted away from basic research over the past decade. But the laboratories remain almost without equal in the industrial sector for the amount of good science they produce. With some of the best scientists accepting voluntary redundancy, and older, senior managers who have supported science at AT&T — "the guys who are protecting us" as one scientist describes them — being strongly encouraged to join the exodus, that may not stay true for long.

Colin Macilwain

Cancer charity spending comes under fire

Paris. A question mark is hanging over the future of France's largest medical charity, L'Association pour la recherche sur le cancer (ARC), following the leak of a report from the Cour des comptes — the government's national audit office — raising serious doubts about the way the charity is run.

ARC's income in 1993, totalling FF581.2 million (US\$118 million) was around five times the combined spending (excluding salaries) on cancer research by the Centre National de la Recherche Scientifique (CNRS) and the biomedical research agency INSERM over the same period. But according to the audit report, the charity's spending on research that year amounted to only FF124.57 million — just one-quarter of its total expenditure.

The charity had previously claimed that three-quarters of its spending in 1993 was on its stated aims of "research, prevention and information". But the report, details of which were published last week by the newspaper *Libération*, also revealed that ARC's accounts put three-quarters of the costs of a fund-raising campaign into this category.

Furthermore, two-thirds of ARC's 1993 spending went to four companies, all belonging to the same group, according to the audit. Contracts with ARC, it claims, accounted for most of the turnover of the companies, while a director of one of the companies is a former director general of ARC, and was assistant director of ARC's magazine *Fondamental* until mid-1994.

The apparent implication is that much of the money collected by ARC was used to finance the development of these companies, although the charity has denied any knowledge of the links between the companies. The audit also found evidence — which it has passed on to the public prosecutor of Paris — of what it claims to be substantial overcharging and excessive commissions by the companies, which were paid by ARC to organize its public relations and produce its two magazines.

The handling of the funds that did go to research is also strongly criticized in the audit. Less than 60 per cent of ARC's grants were approved by a scientific committee, the remainder being passed either by ARC's executive board directly, or at the personal discretion of ARC's founder and chairman, Jacques Crozemarie, who operated an 'emergency' research fund of FF20.17 million in 1993 — a quarter of the sum that passed through scientific committees.

Many grant recipients were also unable to provide the audit with receipts confirming how they had spent their ARC grants. Such lax accounting would have been avoided, according to one INSERM official, if researchers had respected an obligation on scientists working for public research organizations to have the use of grants received

from outside sources scrutinized by the agency itself. But both ARC and grant recipients failed to inform the research organizations systematically of their grants.

One result is that the agencies have lacked precise information on their total budgets. The CNRS and INSERM had estimated that they received FF27.8 million and FF13.9 million respectively from ARC in 1993. But the audit put the real figures at FF53.7 million and FF21.5 million, adding that if such information had been provided by the ARC, it would have led to a better overall use of resources.

The audit's findings are not totally unexpected. Rumours about the running of ARC have circulated for at least 15 years. But the charity — which has enjoyed the support of

many prominent cancer researchers — has fiercely resisted external scrutiny. An investigation in 1990 by the Inspection générale des affaires sociales (IGAS), for example, was stopped after just two months when Crozemarie won a

Crozemarie: questions raised over leadership.

court ruling that, although IGAS is responsible for investigating public authorities, it has no right to investigate charities.

The preliminary findings of the IGAS investigation — which were leaked to the press in 1994 — were almost identical to those of the audit commission (see *Nature* 372, 493; 1994). But ARC, which hotly contested the findings, weathered that storm partly by soliciting letters of support from more than a thousand researchers.

The audit commission was able to investigate ARC only because of a new law passed in 1991, specifically prompted by the government's frustration with its inability to investigate ARC, that gave it the power to do so. ARC's first public reaction was to challenge the audit commission's findings. In particular, Olivier Metzner, a lawyer representing the cancer charity, claimed that half of the charity's spending went on research in 1993, and not 27.2 per cent as the audit had calculated.

But at an emergency meeting of the executive board of ARC the following day, this argument was rejected as indefensible, according to one official present. Instead, the board appointed a six-man working group made up of members of the board — but excluding Crozemarie — to investigate and respond publicly to the audit's allegations, and to propose changes in the way the charity is run, in order to ensure greater transparency in its activities.

But observers have been quick to point out that, by law, anybody subject to an audit investigation has two months to respond, and that both the report and the responses are then made public. Also, some of its members had seen copies of the report several weeks before the so-called 'emergency' meeting.

Furthermore, although the working group excludes Crozemarie, it is nonetheless made up of members of ARC's board, which has endorsed the charity's practices in the past and defended them against similar allegations. Critics argue that the board should have set up an independent committee to review the allegations, suspending its own activities until such investigations have been completed.

Indeed, some claim that the composition of ARC's board has been part of the problem. The boards of many charities are made up of independent individuals responsible for overseeing the running of the organization. But many members of ARC's board receive funding from the charity. One member, for example, Dominique Bellet, director of the immunochemistry laboratory at the Institut Villejuif, near Paris, received the single largest ARC grant in 1993, FF8 million.

"Such a situation hardly leaves the board's members in a good position to exercise control," says one observer. Most of ARC's board members are said to have been personally chosen by Crozemarie. Moreover, according to one board member, the representatives of government ministries sitting on the board did not properly exercise control. The ministries, he claims, often either failed to send representatives to meetings, or sent inexperienced officials who were easily dominated by powerful board members.

One board member, Léon Schwartzenberg, has publicly admitted that the board was lax. The most important consideration, he said, was that income kept increasing. But the audit's findings seem to have shaken the board into action. Pierre Tambourin, the CNRS representative on the board, says that at last week's meeting he had the "impression of taking part in a real board meeting in which influential people were taking part and during which the representatives of the public authorities used their full weight".

Similarly, the working group, meeting for their first time last weekend, suspended all new contracts with the companies accused of overcharging in the audit report, and said that the charity would sign the Charter of Deontology, a code of conduct that has already been adopted by several other charities, including the French Muscular Dystrophy Association.

The board's belated efforts to restore its credibility have included distancing itself ►

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

UK geographers vote to cut links with Shell

London. Members of Britain's Royal Geographical Society (RGS) have voted in favour of dropping the Shell oil company as one of the society's four corporate patrons in protest at the multinational's environmental record in Nigeria. The size of the vote — 204 in favour, with 10 against — has led to calls for the RGS's ruling council to sever the society's links with Shell immediately.

David Gilbert, a lecturer in geography at Royal Holloway College, University of London, who moved the motion during a meeting at the annual conference of the society last week, said in a statement that Shell's environmental and political record in Nigeria "makes the company unfit to be patron of any society representing practising geographers". A spokesman for the company responded by saying that it would be "sad" if the link were severed.

Shell has been heavily criticized for failing to intervene to stop the execution by the Nigerian government last November of Ken Saro-Wiwa and eight other activists belonging to the Movement for the Survival of the Ogoni People (MOSOP). MOSOP claims that oil exploration and drilling from Shell installations in the Ogoni region has devastated the local environment, and that the Ogoni people have been denied their share of the wealth generated by the oil industry.

Gilbert says the RGS's link with Shell compromises the integrity of academic studies of the Ogoni issue, and claims that "Shell is attempting to buy legitimacy in these areas through its patronage of the society".

The charges are denied by Shell. While admitting the existence of "environmental problems" in the area, company officials say they cannot become involved in the domestic politics of a sovereign state. The company also says that it contributed \$25 million towards environmental and community projects in Nigeria last year.

Eric Nickson, a spokesman for Shell International, adds that Shell spent more than £10 million on conservation, environ-

ment and development last year. The company's £40,000 (\$60,000) annual grant to the RGS, he says, helps to fund the Expedition Advisory Centre, "an effective means to help young people to take an interest in environmental development issues".

The vote, which was taken on 5 January at an open meeting at the RGS conference at the University of Strathclyde, Scotland, has no constitutional status within the RGS.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Saro-Wiwa: execution prompted protests.

But the strength of feeling among members at the meeting has prompted the society's governing body to take steps to address the issues raised.

The implications of corporate patronage for a learned society will be the subject of an internal review chaired by Sir Crispin Tickell, a vice president of the RGS and Warden of Green College, Oxford. The names of the society's four corporate patrons — which also include British Airways, Land Rover, and Hong Kong and Shanghai Bank Holdings — appear on the society's stationery.

Meanwhile, the question of the environmental impact of multinational corporations working in the developing world is to be considered at a special RGS symposium later in the year.

According to Tim Unwin, honorary secretary of the RGS and reader in geography at Royal Holloway College, University of London, both steps have been taken in order to show that "a major learned

The working group admits that its aim is to re-establish public confidence in ARC and ensure the charity's survival. This goal is shared by many researchers who argue that funding from charities is essential to the livelihood of their laboratories. Most public funding goes on salaries and overheads, and private income provides much needed funds for equipment and supplies.

Other researchers argue, however, that the ARC affair reveals that the influence of charities is disproportionate compared with the state's total spending on biomedical research, and carries the risk of concentrating control over research excessively in the hands of a few powerful individuals.

Declan Butler

society can be completely open about a controversial subject and move forward in a positive way".

Despite last week's vote and the impending sponsorship review, some observers believe that the 166-year-old RGS is unlikely to make major changes in its present policy of actively encouraging government and industrial sponsorship for its activities, particularly its showpiece multidisciplinary expeditions overseas.

The RGS's political and industrial connections, as well as its access to many diverse sources of funding, are envied by many organizations. Indeed, this is believed to be a key factor behind a vote last year by members of the more academic but relatively impoverished Institute of British Geographers to rejoin the RGS after breaking away in 1933.

Other learned bodies, such as the Royal Society and the Institute of Physics, also maintain links with industry — but have no tradition of corporate patronage. The oil company Esso sponsors an award and a medal in conjunction with the Royal Society. The 21,000-member Institute of Physics (IOP) operates a scheme in which commercial organizations can become 'corporate affiliates' by paying £2,000.

Asked if the IOP would consider an application for affiliation from Shell, Susan Partridge, industrial affairs manager at the IOP, says such a request would not be automatically ruled out. "Affiliates are not screened on commercial criteria, as that would exclude important branches of physics," she says. "Physics comes first."

Environmental organizations have also found ways to harness corporate support. Ten per cent (£2 million) of the income for the World Wide Fund for Nature (WWF), for example, comes from corporate activity, including a scheme in which the WWF licenses its panda logo for products that are produced using eco-friendly methods.

WWF has also accepted both unsolicited as well as project-oriented donations, known as Corporate Partnerships, from Shell. The former includes a sum of £9,600 received between 1992 and 1995. A Shell-supported project on wildlife habitat and forestry plantations is an example of the latter.

Robin Pellew, chief executive of WWF, says candidates for Corporate Partnerships are screened for ethical and environmental credibility. "WWF did not know of Shell's activities in Nigeria" when the company was last approached. WWF, he adds would not approach Shell in the future, but would not turn down unsolicited donations.

A spokesman for Shell, however, says that its £9,600 donation to WWF was not unsolicited. "A number of ideas were circulated by WWF and we made a donation to one of them."

Ehsan Masood

► (from page 103) from Crozemarie. Responding to public statements by Crozemarie that he would dissolve the charity rather than resign, the working group last weekend reminded him that such a step would require approval of ARC's general assembly. "It is divorce," said one board member.

According to an official present at the emergency meeting, most board members also opposed keeping Crozemarie as chairman of the charity. The only reason Crozemarie was not asked to resign immediately, he says, is that the board wanted time to find an "irreproachable" successor, such as a retired judge or a leading scientist from a field other than cancer research.

Dominance of older scientists 'threatens research' in India

New Delhi. The average age of scientists working for government scientific organizations in India is increasing at an alarming rate, according to a study carried out by the Department of Science and Technology (DST), which says that this should be a "matter of concern" to the government.

In particular, the study has found that about 80 per cent of scientists occupying senior positions are over 50 years of age. One-third of those in junior levels, and two-thirds of middle-level scientists, are also over 50 years old.

The study was carried out by A. R. Rajeswari of DST, and covered researchers in 10 agencies operating 247 research and development (R&D) laboratories, employing 31,340 scientists. It found that almost half of the researchers joined their organizations before 1975, and pointed out that they "need retraining to keep abreast of the latest developments in science and technology".

Unless action is taken, says the study, leaving the work of these institutions in the hands of older scientists may reduce the overall viability of Indian research. It says that the current situation is due to rigid employment structures and regulations governing appointments and promotions. For instance, "senior positions were invariably reached only by seniority measured in sheer number of years in an organization, rather than by merit promotions".

Sri Krishna Joshi, who recently retired as head of the Council of Scientific and Industrial Research (CSIR), admits that the ageing of research personnel "has adversely affected the innovativeness and creativity of CSIR".

Joshi claims that the average age of CSIR scientists is "rather high" because a ban on recruitment has been in place for the past decade, as a result of a lack of resources. The ban was lifted only a few months ago.

The DST study has also highlighted the lopsided gender profile of India's scientific agencies. Of the 301 senior positions analysed, only two per cent were occupied by women.

While most male scientists belonged to the older age bracket, the majority of female scientists were younger (under 43), apparently because female entry into science is relatively recent.

The study also shows that India's women scientists may be more productive than their male counterparts: on average more patents are granted to female than to male researchers.

K. S. Jayaraman

Money woes fail to weaken Japan's support for science

Tokyo. Despite a continuing recession, as well as the need for massive government payments to bail out failing banks and housing and loan corporations, Japan's science-related ministries and agencies have managed to win surprisingly large increases in the budget for the 1996 fiscal year, which begins on 1 April.

Impoverished university researchers will fare particularly well. Not only is there a large increase in the grant budget of the Ministry of Education, Science, Sports and Culture (Monbusho), but there are also new grant schemes open to university researchers from the Ministry of International Trade and Industry (MITI) and the Science and Technology Agency (STA), marking a new era of cooperation between these agencies.

The total government budget of ¥75,105 billion (US\$750 billion), an increase of 5.8 per cent over last year, was approved by the cabinet at the end of last month.

The budget must still be formally approved by the Diet, Japan's parliament. But science-related elements are rarely, if ever, changed in the approval process, although a controversial outlay of ¥685 billion to bail out failing financial institutions may well delay approval by the Diet — and could even bring about a general election.

The comparatively large increases for science agencies, from 5 to 10 per cent, stand in sharp contrast to the fate of other government departments. The increase for overseas development assistance, for exam-

ple, will be only 3.5 per cent, and for defence only 2.6 per cent. Both have for many years tended to receive the largest increases.

Science and technology are now in favour, thanks to lobbying by a small group of comparatively young and powerful politicians from various parties who passed a law late last year requiring the government to do more to support basic research (see *Nature* 378, 227: 1995). The law does not commit the government to any specific increases. But it has given the science-related ministries and agencies additional leverage in their budget negotiations with the Ministry of Finance.

Monbusho's competitive grant budget for university researchers has for the first time exceeded ¥100 billion (US\$1 billion), representing an increase of more than 10 per cent, and in doing so, achieving an ambitious goal set by the ministry in 1992 when the grant budget stood at only ¥65 billion.

Also notable is a new grant scheme for 'strategic' research from the STA. With a budget of ¥15 billion in its first year, this will provide five-year grants ranging in value from a few tens of millions of yen to ¥200 million a year, and will be open to university researchers, as well as researchers in national research institutes. Similarly, MITI has a new programme for 'creative' research, with a budget of ¥2.6 billion.

These grant schemes follow on from programmes introduced by MITI and STA in one-off supplementary budgets last year (see *Nature* 377, 378: 1995). They represent the

first time government agencies other than Monbusho have been able to provide large grants to university researchers. STA's success in winning the full ¥15 billion requested from the finance ministry is particularly remarkable; normally, such budget requests are trimmed in the end-of-year negotiations, and it is a sign of the new climate that STA obtained the full sum it was seeking.

In line with a recent call by the government to double the number of postdoctoral fellows in Japan to 10,000 by the end of this decade, Monbusho and STA have each won large increases for their postdoctoral fellowship schemes. Monbusho has also been awarded a substantial increase for its new Centre of Excellence (COE) programme, which provides large sums of money, including funds for postdoctoral fellows, for a few select groups of university researchers.

David Swinbanks

HIGHLIGHTS OF JAPAN'S BUDGET FOR SCIENCE FOR 1996 (in billion yen: US\$1 = 100 yen)

		% Change
Science and Technology Agency		
Total R & D budget	692.8	+7.2
Special promotion funds	21.5	+16.2
Space	177.9	+5.2
Nuclear power	319.5	+3.5
Spring-8	16.7	+10.6
Human Genome	2.4	+4.1
Human Frontier Science Programme	3.6*	+1.7
Strategic basic research	15.0	New
Ministry of International Trade and Industry		
Total R & D budget	319.0	+5.3
Agency of Industrial Science and Technology	144.0	+7.4
Industrial scientific technology	26.4	+6.3
Creative R & D projects	2.6	New
Ministry of Education, Science, Sports and Culture		
Grants in Aid of Research	101.8	+10.2
COE Programme	11.4	+23.8
Postdoctoral fellowships	11.2	+33.5

*Includes 1.4 billion yen from MITI.

Russians wrangle over legal status of academy property

Moscow. The Russian President, Boris Yeltsin, has written to Vladimir Shumeyko, the Speaker of the Federal Council (the upper chamber of the Russian parliament) complaining that the council has altered the text of a planned law on science and science policy by the "deliberate substitution of some pages" sent to the president for his approval.

Yeltsin is also said to have informed Russia's newly appointed procurator-general, Valery Skuratov, that the alterations may have infringed criminal laws covering the limits of the authority that can be exercised by public officials.

In his letter to Shumeyko on 23 December, Yeltsin says that analysis of the text of the law approved by the Federal Council earlier in December revealed seven alterations from a text earlier adopted by the State Duma. Under the Russian constitution, the Senate has no power to change anything in the laws adopted by the lower chamber. The apparent differences between the two versions have led Yeltsin to demand that the original version — and not the amended version — be sent to him.

That has raised some eyebrows in Moscow, as it is not unusual for the Senate to edit the text of documents passed by the Duma before officially approving them. Some scientists are linking Yeltsin's letter with his recent decision to appoint Yuri Osipov, the president of the Russian Academy of Sciences (RAS), to Russia's ruling body, the Presidium.

This act was seen as largely symbolic — reflecting the president's desire to demonstrate his support for science — and the letter to Shumeyko is being viewed by some as similarly motivated. But the weakness in this explanation is that the amended version of the science law would put the academy in a stronger position than the original version by giving it ownership of its own facilities.

According to Vladimir Grachev, chairman of the Senate Committee on Science, Culture and Education, when the bill was received from the State Duma on 30 November, members of his committee found "some drawbacks of mainly stylistic origin" and asked Duma officers to make corrections on two pages.

The following week, the first deputy speaker of the Duma, Michail Mityukov, sent to the Senate an official letter asking for seven pages to be replaced by a new version "due to technical mistakes". These appear to be the changes referred to by Yeltsin in his letter.

The conflict appears to have arisen over an additional revised page, sent personally to the Senate by the speaker of the Duma,

Ivan Rybkin, and, according to the Senate's legal department, containing changes that would alter the main thrust of the new law.

The original text had specified that the RAS and branch academies were to be given the rights to "control their activities and federal property" and to "possess, utilize and command the federal property *given to them in administrative control and economic management*" [emphasis added].

The proposed revision, however, states that the academies should have the "ownership" of all their property, including not only the scientific equipment but also the buildings, dwellings, sanatoriums, kindergartens, hospitals and so on. Shumeyko, the speaker of the council, subsequently sent a 70-page letter to Skuratov, suggesting that the changes in the bill spotted by Yeltsin had been made by the lower, and not the upper, chamber of the Russian parliament.

Asked how this could have happened, Nikolai Vorontsov, chairman of the former Duma's science subcommittee, who had presented the science law to a plenary debate, claimed that Viktor Shevelukha, a member of the communist faction, had managed when acting as chairman of the full Committee on Culture, Education and Science, to get Rybkin's signature on the letter to the council asking for reference to "control" to be changed to "ownership".

Interestingly, the version of the bill distributed during a meeting on science legislation in the State Duma last year (see *Nature* 378, 4; 1995) included a footnote indicating that Shevelukha was proposing to add the words "or in ownership" in the relevant clause. When the State Duma later rejected this proposal, Shevelukha managed to insert it in the text of the law himself — omitting the word "or".

Meanwhile, despite widespread discussion of the need to support Russian science during last month's election campaign for the Duma, little mention of the topic has been made since the victory of the newly resurgent communist forces.

Vorontsov was one of few prepared to comment. "No doubt, the things will change to the worse", he says, arguing that the science most likely to receive support is that related to the needs of government agencies — and primarily military research — "but not fundamental science".

Since Vorontsov, a strong supporter of basic science, failed in his bid for re-election to the Duma, his forecast may well come true. But as the proposed science law has not yet been signed by Yeltsin in its present pro-communist version, there is still some hope that basic science will not be totally neglected.

Carl Levitin

Cancer fund likely to boost basic research under new director

London. Paul Nurse, currently director of laboratory research at the Imperial Cancer Research Fund (ICRF) in London, is to take over as director general of the charity on 1 September. He will succeed Sir Walter Bodmer, who has been director general since 1991 of a charity which employs over 1,000 people and spends in excess of £50 million a year on cancer research.

Bodmer is to become principal of Hertford College at the University of Oxford. Nurse, who is 46, headed a laboratory at the ICRF between 1984 and 1987. He returned to the ICRF in 1993 after six years as Iveagh Professor of Microbiology at the University of Oxford.

The news of his appointment has been welcomed by ICRF researchers, who have faced six months of uncertainty following Bodmer's announcement of his departure last July. "Paul is a very highly regarded scientist," says one ICRF biochemist. "His goal is the pursuit of excellence. It might make people feel a little insecure, but you don't get a first class institution for nothing."

Nurse, whose research interests focus on an understanding of the genes which prompt cells to multiply, says one clear change of emphasis from his predecessor will be his desire to protect and promote basic research. "My predecessor thought that we understood enough basic research [about cancer]," says Nurse. "My view is that our knowledge remains limited, and that basic research must continue."

Nurse says another priority will be "translational research", involving the effective transfer of a laboratory idea into the clinical arena, and getting clinicians talking more productively with scientists.

Nurse has a wide range of extra-laboratory interests. For example, he has been active in promoting the public understanding of science, arguing that more emphasis needs to be placed on communicating the process of science, rather than publicizing its "gizmo" applications.

He is also a committed member of the opposition Labour Party, and has little hesitation in voicing his opinion that science has not prospered under the present government. "We need a change of government to allow us to have a fresh look at how science can be supported in the United Kingdom," he says. Ehsan Masood

Seismologists prepare to reap test ban data

Washington. Seismologists are set to win unrestricted access to a vast amount of data that will be gathered by a new international seismic monitoring system being established to listen out for nuclear tests under the proposed Comprehensive Test Ban Treaty (CTBT).

Proponents of the treaty hope that it will be agreed this summer. But even if its international ratification takes years, earthquake scientists and others stand to benefit immediately from the plans for open access. Data gathered by a prototype monitoring system from fifty stations worldwide has already been released on the World Wide Web by the Air Force Technical Applications Centre (AFTAC), the US agency with responsibility for nuclear test monitoring.

A panel of the National Research Council, the research arm of the National Academy of Sciences, reported last

November on what was needed to ensure that the treaty's monitoring system will generate the widest possible scientific benefits. Its report, *Seismological Research Requirements for a Comprehensive Test Ban Monitoring System*, says it will be easy to widen the bandwidth of vibration monitored by the system to include the very low frequencies — down to 0.003 Hz — of interest to earthquake scientists, as well as the higher frequencies which betray underground nuclear explosions.

The report also called for data from the system to be presented in the Standard for Exchange of Earthquake Data (SEED) format. Thorne Lay of the University of California at Santa Cruz, who chaired the panel, says that many of its recommendations have already been incorporated in the CTBT treaty. Proponents hope the treaty will be signed this summer, and even if ratification is delayed, construction of the

monitoring system will begin at that point.

"This is a sea-change, because work that has been classified in the past is going to produce data for the research community around the world," says Lay. The data will not only be twice as voluminous as that currently available, he adds, but will also be available much more quickly to scientists.

The panel also found that an extensive programme of both basic and applied research would be needed to meet the US policy goal of detecting nuclear blasts of as little as 1 kiloton, even if ignited in a cavern to reduce their seismic impact. Such a blast will produce a disturbance of 2.5 on the Richter scale, while existing technology could only identify an explosion of 4.5, or 100 times the magnitude.

The United States intends to continue secret monitoring for nuclear blasts separately from the proposed international effort.

Colin Macilwain

Change of diet unsettles nutrition research in Germany

Munich. A DM200-million (US\$140-million) research building owned by Germany's Ministry of Agriculture, and designed to house 600 nutrition researchers, could become a white elephant when completed in 1997 because of a major reorganization of nutrition research in Germany, coupled with sharp funding cuts.

Reflecting pressure on the government to cut public spending and increase the effectiveness of its research budget, the ministry announced plans last August to cut 860 of its 2,600 positions in its five nutrition research centres by the year 2005. The main opposition, the Social Democratic Party, is now demanding a parliamentary debate, claiming that uncertainty about the ministry's plans for the centres is causing alarm among staff.

The ministry wants to achieve the cuts partly by closing the federal meat research institute (BAFF) in Kulmbach and the equivalent body for corn, potato and fat research (BAGKF) in Lower Saxony. Some of the research from both would be transferred to two broad-based research institutes, in Karlsruhe and Kiel, while the overall target for cuts would be met by placing a ban on new staff at the remaining three centres.

Following strong opposition from the research centres themselves, as well as from the state of Bavaria, in which the BAFF is located, the ministry is now revising its initial proposal. Both the federal government and the researchers involved say that they want a quick decision. But the ministry argues that, as the final plan is tied in with the federal budget proposals for 1997, no decision can be made for several months at the earliest.

The reorganization plans stem from a

desire not only to save money but also to deal with long-standing complaints about the inefficiency of the organization of nutrition research in Germany. According to Reinhold Dörre, for example, director of an umbrella organization of scientific societies concerned with agricultural research, much of the work in government laboratories duplicates that carried out elsewhere, such as in universities, while the idea that a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Robert Haring

Feeling the heat? Nutritional research faces staff cuts and other changes in Germany.

research centre should focus on a single food product, such as meat, is outdated.

Ten years ago, already sensitive to such arguments, the ministry endorsed a plan to expand the broader-based Federal Nutrition Research Centre (BFE) in Karlsruhe by expanding its role as an interdisciplinary nutrition research centre, increasing the staff from 250 to 600 and providing a new building for them.

But as the building nears completion, a large question mark stands over its future, partly as a result of further shifts in government policy that have given the Karlsruhe centre a less dominant role in practice. As a

result, far from growing in the past decade, staff numbers at the BFE have gradually fallen to 200. "It is all so uncertain", says its director, Ulrich Spieß. "No-one knows what will happen when. And we will be lucky to be left with 160 staff after the cuts."

Spieß would like to share the new building with local university departments in related subjects, bringing the spin-off of mixing basic with applied research. But the ministry is holding its cards close to its chest. A spokesman, Ulrich Metzner, says that the ministry has not made any decision on the building's future, although it is keen not to lose its investment, making it "unlikely that it will agree to share with [any] university".

Meanwhile, despite behind-the-scenes support from Bavaria's prime minister Edmund Stoiber, pressure on agricultural research to rationalize its approach and increase its efficiency means that hopes among the staff of the BAFF and the BAGKF that closure of their institutes can be avoided are probably misplaced.

Indeed, even the Social Democrats, in their demands for a parliamentary debate, are being careful to emphasize that they are seeking an end to uncertainty rather than a reprieve for the threatened research centres.

Furthermore, the federal research minister, Jürgen Rüttgers, told the German Society for Agriculture and Environment Politics last year that agricultural research receives a disproportionately large amount of public research funds. Agriculture contributes less than one per cent of Germany's gross national product, but 2.5 per cent of German research spending goes to agricultural research, he said.

Quirin Schiermeier & Alison Abbott

Purpose and function of IPCC

SIR — Your leading article, “Global warming rows” (*Nature* 378, 322; 1995) displays a misunderstanding of the purpose and functioning of the Intergovernmental Panel on Climate Change (IPCC) and its three working groups. You characterize these three groups as dealing respectively with “questions of climatic evidence, the consequences of warming, and with ‘advice to policymakers’”, and you write as though homogenous, closed groups of experts write the reports and the summaries. This is mostly incorrect.

The IPCC is the most extensive and carefully constructed intergovernmental advisory process ever established in international relations. Governments nominate experts — people with proven academic track records — that then form writing teams selected for breadth of expertise and geographical balance. The three working groups are not in any way constructed to give a linear path from science to impacts to policy advice. Rather, they are concerned with assessing the literature on (1) climate science, (2) technical aspects of impacts and response options, and (3) economic and other cross-cutting issues. The teams write the main chapters (and chapter summaries), which generally go through widespread peer review — with hundreds of people and groups worldwide invited to comment. The content remains under the control of the writing teams.

The ‘Policymakers’ Summary’ is not a technical summary. Although usually drafted by the experts to put all key findings ‘on the table’, the final text represents a painstakingly negotiated statement of what governments officially accept as a balanced account of the state of knowledge and reasoned judgement, based on the chapters. Governments cannot alter the chapters, and authors cannot alter the Policymakers’ Summary. In other fora (the UN Climate Convention) governments will debate appropriate policy responses, and obviously may wield the words they have agreed in the IPCC Policymakers’ Summary as ones carrying particular force. But the IPCC is precluded from making any policy recommendations, and as a body does not.

The biggest threat to rational action on climate change is ignorance on the part of governments — ignorance that would find expression in governments choosing their preferred facts and preferred experts to suit national interests or particular pressure groups. The IPCC forms a critical, mutually agreed process to establish a basis of internationally accepted knowledge, and without such a process the negotiations would be rudderless and irrational. And the process of developing IPCC reports plays at least two further roles: it

educates government delegates about the findings of research; and it stimulates critical debate between research communities that might otherwise try and ignore each other’s existence.

Your suggestion that Working Groups II and III be abandoned is extraordinary. Do you really think that negotiations would proceed better if there were no such assessment of literature on the impacts of climate change, the technical options for mitigation, or of the economic and related analyses surrounding the issue? As it happens, I have sympathy with some of the reservations expressed by governments about some of the literature on valuation of impacts (and valuation of statistical life). But I would defend absolutely both the right of the chapter authors to say what they think, and the right of governments to voice their reservations and concerns. I have had my share of intellectual clashes on other parts of the Working Group III report and the report is not as any of us would have written it alone; but all those involved — authors, government representatives, reviewers and even perhaps a few media commentators — are wiser for the experience.

Of course the IPCC is not perfect. Everyone works under time pressure. Some writing teams turn out not to encompass all valid views, and some may not do full justice to review comments; experts suffer personal preferences and other human faults. Judgements also have to be made about inclusion of important information that reaches publication only in the later stages of the process. Governments could be given a stronger hand in policing all this — but only at the risk of greater politicization of the reports. Similarly, authors could be given stronger power over the Policymakers’ Summary — but only at the cost of reducing its political authority and relevance as a governmentally agreed document. Like democracy, the IPCC is far from perfect, but I have yet to see a better system of advice in international affairs.

Michael Grubb

(Lead Author, Working Group III)
Energy and Environmental Programme,
Royal Institute of International Affairs,
Chatham House,
10 St James’s Square,
London SW1Y 4LE, UK

SIR — I should like to clarify the situation in response to your reporting on Working Group 3 (WG3) of the IPCC (*Nature* 378, 119; 1995, and subsequent issues).

In the IPCC procedures established by the governments, lead authors prepare technical assessments of published litera-

ture on an agreed range of topics. These reports are subject to peer and government reviews. The government representatives then have an opportunity to select from these technical reports those findings of special relevance to policy-makers, or in some instances to comment on the findings. This results in a Summary for Policy Makers (SPM), which is the responsibility of government representatives to the working group. The technical reports remain the responsibility of the lead authors.

In the case of WG3, at plenary meetings of more than 70 country representatives and a number of non-governmental organizations (NGOs) in Geneva (25–28 July 1995) and Montreal (11–13 October 1995), an SPM was approved line by line, and the technical report, including all chapters, was accepted.

There was some contention, as you have reported, about matters contained in Chapter 6, “Social Costs of Climate Change: Greenhouse Damage and the Benefits of Control”, as well as several other chapters. In the approved SPM, country representatives expressed concern about the way in which the economics literature treats non-market damages, especially the risk of increased mortality (value of a statistical life). In the SPM they make it clear that “while some regard monetary valuation of such (non-market) impacts as essential to sound decision making, others reject monetary valuation of some impacts, such as risk of human mortality, on ethical grounds”. The SPM goes on to express the view that “societies may want to preserve human life in an equal way”.

The corresponding technical chapter by the lead authors draws a distinction between a “prescriptive approach” to the value of a statistical life and a “descriptive approach” (Box 6.1). A prescriptive approach asks the question, “at how much ought a statistical life to be valued according to ethical or other criteria?” The authors conclude that “[a]n obvious implication of this approach is that all lives will be treated equally”. The text then goes on to outline the “descriptive approach” by which most of the economics literature (in many fields, not just climate change) evaluates the costs that society is willing to incur to protect lives, or “the willingness to pay”. Thus, the “ethical” or “prescriptive” approach preferred by country delegates is explicitly recognized in Chapter 6, although the literature available for assessment was based on a descriptive approach. In addition, both Chapter 6 and the SPM make the point that: “For individual nations, or if alternative assumptions are made about the value of a statistical life (damage figures) could be much higher” (wording from Chapter 6).

Chapter 6 contains a great deal of valuable information to help policy-makers, especially those in developing countries. It provides estimates that show climate change damages in developing countries are likely to be up to 9 times greater, proportionately, than damages to developed countries. It also shows that the net aggregate effects of projected climate change would probably be global damages of the order of a few per cent of gross domestic product. This gives an economic rationale for acting to reduce greenhouse gases that goes beyond energy efficiency and other measures that are worth doing for their own sake. These and other findings in the chapter, and in the balance of the WG3 report, support actions both on greenhouse gas reductions and to assist developing countries. The SPM and chapters have already been cited favourably by the Climate Action Network (Environmental NGOs) in their submission to the Ad Hoc Group on the Berlin Mandate of the Conference of Parties to the Framework Convention on Climate Change.

The final Plenary Session of the IPCC in Rome (11–15 December 1995), with some 120 countries represented, has also accepted unanimously the Summary for Policy Makers and all chapters of WG3 (as well as those of Working Groups 1 and 2). This acceptance was a recognition that, while government representatives may not necessarily agree with all statements in all the authored chapters, they find a wealth of information and valuable assessments in the Technical Chapters of WG3 that will assist future climate negotiations.

James P. Bruce

(Co-Chair IFOC, WG3)
1875 Juno Avenue,
Ottawa,
Canada K1H 6S6

SIR — It was reported in a recent News story (*Nature* 378, 329; 1995) that I declined a request from Pat Michaels for detailed output from a climate model, the details of which are cited in the second scientific assessment of the IPCC. He claimed that he needed the gridpoint data to review the IPCC report. At the time, I sent Michaels the current draft of the paper cited in the IPCC report (and subsequently published in *Nature* 376, 501; 1995). Michaels had expressed concern that scattering from anthropogenic sulphate aerosols would not diminish the expected warming over the Arctic due to increasing greenhouse gases. So I also sent another preprint, subsequently published in the *Journal of Climate* (8, 2364–2386; 1995) which explains how the Arctic can be cooled by aerosols, even if they are not located over the Arctic. In accordance with normal practice, the raw data are not normally released until the results are formally published. Hadley Centre model data are the property of the UK Department

of the Environment and not IPCC, and are made freely available to bona fide researchers in due course through the UK Climate Impacts LINK Project.

John F. B. Mitchell

Hadley Centre for Prediction
and Research,
Meteorological Office,
Bracknell RG12 2SZ, UK

SIR — Aubrey Meyer *et al.* assert (*Nature* 378, 433; 1995) that a chapter should be excluded from the Second Assessment Report (SAR) of the IPCC. The chapter concerns “Social Costs” of climate change and has been prepared by IPCC Working Group 3 (WG3).

This SAR Summary has been approved but the chapter prepared by WG3 disagrees with the summary. Meyer *et al.* observe that the IPCC procedures do not permit amendment to approved summaries, and they do not like the WG3 chapter. For these reasons, they assert that the chapter should be excluded from the SAR. But the correct amendment would be a change to IPCC procedures, not a change to the SAR. The IPCC summaries should be written and approved only after the work they summarize has been completed.

The IPCC approved the summary of its 1994 Scientific Assessment before that report was written. This resulted in misleading data being ‘tailored’ to fit the summary in that report. Now, either the summary of the SAR will ‘summarize’ a chapter that is not in the SAR, or the summary will say the opposite of a chapter that it claims to summarize. Neither of these options should be acceptable to scientists of integrity.

Do the signatories of the letter from Meyer *et al.* never conduct peer review of other people’s work? Would any of them recommend publication of a paper that contained a ‘summary’ that ‘summarized’ work not in the paper? And would they recommend publication of a paper that contained a ‘summary’ stating the opposite of conclusions in the paper?

The IPCC purports to be a scientific organization providing scientific advice to politicians. Many observers have suspected that it is a primarily political organization. The letter from Meyer *et al.* is all the confirmation they require of their suspicion.

Richard S. Courtney

31 Rivelands Road,
Swindon Village,
Cheltenham,
Gloucestershire GL51 9RF, UK

Correspondence

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to nature@nature.com

Two sides of Spanish science

SIR — In the past decade, *Nature* has carried many articles about science in Spain, but it has not previously been remarked that there are two kinds of researchers in Spain. One, which could be called ‘Science Citation Index (SCI) scientists’, holds the real power in the Spanish scientific community.

These SCI scientists work in such fields as biochemistry, biology, biophysics, chemistry, cytology and histology, endocrinology, immunology, materials science, medicine, neurosciences, optics, physics and spectroscopy. These subjects have in common the ease with which they can be published in SCI ascribed journals because of their widespread interest. SCI scientists do not carry out much field work, they frequently have contacts with overseas researchers and they carry out few educational activities.

The other group, which might be called ‘parallel Spanish Scientists’, work in areas that are not well developed in SCI journals. They often carry out field work in such areas as agriculture, ecology, engineering, forestry, food science, geography, geology, geosciences, oceanography, ornithology, botany, veterinary sciences and zoology. These parallel Spanish scientists usually work in local areas and publish in serious Spanish journals that are not in the SCI lists. They often know French or German (rather than English) and carry out other activities such as teaching, writing books, organizing meetings or running laboratories.

The easy computerized access to the SCI lists to measure productivity has simplified the government’s task in awarding grants, increasing salaries, allocating projects, forming boards of examiners, providing posts for civil servants and so on. SCI productivity is the crucial number that governs the life of all Spanish researchers. This pragmatic transformation of Spanish science, which was accelerated in 1986 when Spain joined what is now the European Union, has two different results: bolstering scientists working in exportable subjects who are obviously happy with this policy, and causing great discontentment among the ‘parallel researchers’ working in regional and local subjects. In spite of the apparent unfairness of this system, many ‘parallel Spanish scientists’ are working increasingly on papers included in SCI journals. Meanwhile, we all watch with sadness the decline in quality of contributions to Spanish written journals.

Javier Garcia-Guinea

Departamento de Geología,
Museo Nacional de Ciencias Naturales,
c/ Jose Gutierrez Abascal 2,
28006 Madrid, Spain

Trusting in traditional cures

SIR — At a recent forum, some distinguished scientists saw the rising popularity of alternative medicine as another indication of irrational thought, and the US National Institutes of Health (NIH) Office of Alternative Medicine as symptomatic of the anti-science movement¹.

Although there are many non-scientific elements in traditional and herbal remedies, they have contributed to modern conventional medicine and it would be improper to dismiss them. For example, modern research suggests that traditional Chinese medicine has a scientific basis. In drug discovery, the development of new drugs based on leads from Chinese herbs is more efficient than the conventional method of random screening because it is guided by experience from a long history of clinical practice. In the United States,

several new drugs of Chinese herb origin have recently been commercially developed. These include derivatives from the antimalarial qinghaosu such as Paluther²; huperzine A, a promising drug for Alzheimer's disease³; and trichosanthin (GLQ 223) which has completed its Phase II clinical trials as an anti-HIV therapeutic^{4,5}.

Traditional Chinese medicines have worked effectively in some instances where Western therapies have failed or proved to be insufficient. In a double-blind British clinical trial, a formula of Chinese herbs has produced impressive responses in severe atopic eczema that was resistant to conventional treatment⁶. Another controlled clinical trial conducted in Japan showed that *sho-saiko-to*, a Chinese herbal formula, helps to prevent liver cancer in patients with cirrhosis⁷.

Work on traditional Chinese medicines may also bring refreshing concepts to conventional medicine. For example, an important feature of traditional Chinese medicine is the preference for composite prescriptions (combinations of herbs) over single active substances. This practice is not readily compatible with modern medical theories. However, the clinical trials described in refs 6 and 7 show that herbal combinations sometimes produce synergistic therapeutic effects on patients.

Some diseases are now resistant to con-

ventional treatment and no fundamentally new therapies have recently been developed for any major forms of cancer. In this context, I think scientific research on traditional and herbal remedies is necessary as an alternative approach to the discovery of new and better therapies, and the Office of Alternative Medicine's plan to begin some clinical trials of Chinese herbs⁸ should be implemented.

Tianhan Xue

460 Whitman Street, Suite 72,
Santa Barbara, California 93117, USA

A drop in the ocean

SIR — The discussions in *Nature* about the Brent Spar oil platform can be illuminated by a comparison with what the oceans have borne in the past. The German Navy, for example, sank some 10 million tonnes or more of shipping in the North Atlantic during the Second World War, much of it as oil in tankers and in fuel tanks in merchant and naval vessels. Goodness knows what heavy metals went to the sea floor with the cargoes of military hardware and lubricants. How can the German greens and others panic about 100 tonnes of oil in Brent Spar?

Charles Kittel

PO Box 368,
Philo, California 95466, USA

1. *Nature* **375**, 439 (1995).

2. *Wall Street Journal*, 28 August 1992, B5.

3. Borman, S. *Chem. Engng. News* 20 September 1993, 35–36.

4. Byers, V. S. *et al. Aids Res. hum. Retrovir.* **10**, 413–420 (1994).

5. Byers, V. S. *et al. Aids* **4**, 1189–1196 (1990).

6. Sheehan, M. P. & Atherton, D. J. *Br. J. Dermat.* **126**, 179–184 (1992).

7. Oka, H. *et al. Cancer* **76**, 743–749 (1995).

8. *Nature* **370**, 591 (1994).

Are you still just sitting there?

Of course we don't want you to run chromatography like this. But we do want you to do it this easily. UNICORN™ 2.00 automated control software assures you of an easier transition from one stage of development to the next. A much easier transition.

From the moment you set up your first run on UNICORN 2.00, there's no need to use any other control software. At any stage. Ever.

UNICORN 2.00 not only saves you time—there's only one program to learn—it saves you effort as well. That's because there's no transferring data between your different systems.

Just call us at 1 (800) 526 3593 in the United States, or +46 18 16 5011 from the rest of the world, to find out more about UNICORN 2.00. It truly is the only automated control software you'll ever need.



**Pharmacia
Biotech**

Uppsala, Sweden. (And the rest of the world)

Science's fate in Congress

SIR — I was puzzled by your leading article praising Republican science policy in the US Congress (*Nature* 378, 115; 1995), particularly in describing Bob Walker, the chairman of the House of Representatives Science Committee, as “an effective leader for science in Congress”. My impression is quite the opposite.

Republicans have made funding for research and development a low priority. Their budget would impose research and development cuts estimated at more than 30 per cent over the next seven years (after inflation), even as overall federal spending would remain roughly constant.

The Republican budget also takes an ideological approach to research and development funding, shattering the bipartisan consensus that had evolved over many years. While sparing some areas of basic research, the budget targets a wide range of programmes in energy and environmental research and civilian technology development. I believe that public funding for these useful but non-commercial areas of science and technology is essential for maintaining public support for science and technology in general.

Finally, Walker has not made the Sci-

ence Committee one whose “output counts for something”, but instead has made the committee largely irrelevant. Decisions are made by the Republican leadership and by the Budget Committee, and rubber-stamped by the Science Committee. Committee members, Republican and Democrat alike, have had little substantive influence.

Mark Goodman

7004 Exfair Road,
Bethesda, Maryland 20814, USA

Gene patenting

SIR — I am a molecular biologist, with little legal expertise, who is unhappy about gene patenting. I refer in particular to the Commentary article by George Poste (*Nature* 378, 534; 1995). My main objection is the obvious fundamental point: how can a discovery be redefined as an invention? This issue is presented as a legal *fait accompli* and skipped over by Poste. An explanation offered is that gene discovery requires special skills, but the difficulty of making a particular discovery hardly justifies calling it an invention, especially when the techniques involved are now routine. It has taken evolution 3.5 billion years to produce the DNA sequences that constitute our genes, yet at the stroke of a legal pen these can be deemed as human

“genomic inventions”. A not too dissimilar line of reason would have entitled Isaac Newton to patent gravity.

Fergus Davison

Neuropathology Department,
Institute of Neurology,
London WC1N 3BG, UK

Unthinking plankton

SIR — Although we have come to expect, but not readily tolerate, teleological language in television programmes on wildlife (including those of Sir David Attenborough) and in popular scientific writing, such a style should certainly not appear in *Nature*. Zooplankton do not “migrate to greater depths to reduce (my italics) their chance of being detected by visual predators (fish)” (L. De Meester *et al.* *Nature* 378, 483–485; 1995). Nevertheless, the paper clearly indicates that their behaviour does have that effect, that is, such migration of zooplankton reduces their susceptibility to predation. This is a statement of fact based on observation and experiment, with no implications that zooplankton had great thoughts about their strategy.

D. H. Lewis

Department of Animal & Plant Sciences,
University of Sheffield,
Sheffield S10 2TN, UK

Yep, but I've transferred my
lab methods to the production hall on
this new UNICORN 2.00 software



Basic research fighting for survival

David Swinbanks

A recent conference on "nurturing creativity in research" revealed some extreme examples of the growing tendency of governments to drive research towards strategic socioeconomic goals at the expense of basic research.

Canberra. A conference* in Australia last November, co-sponsored by *Nature* and the Australian National University (ANU), was a rallying point for scientists and science policy-makers to champion basic research at a time when governments around the world are pushing research increasingly towards strategic national goals intended to boost the economy. Although focusing on Australia, the meeting also included speakers from the Asia-Pacific region, Europe, the United States and Canada. But it was Philippa Black, president of the Royal Society of New Zealand, who gave the most stunning example of the harm done to basic research through strategic planning.

New Zealand began restructuring its research system in 1986 in response to the dire economic situation of the country. The discipline-based government research institutes were replaced with ten Crown Research Institutes run as profit-making companies serving productive sectors such as the wool industry or natural resources. Some, such as the Institute of Environmental and Scientific Research, which sells forensic services, have earned substantial commercial revenue. But others have not fared well, and the smallest failed in 1994.

Parallel changes were made in the research grant system. Unlike research councils or foundations elsewhere, New Zealand's new 'Public Good Science Fund' — the only source of government non-operational research funds — is a truly 'strategic' fund with allocations based on 'science outputs'. It started with 40 categories covering such things as 'sheep production' and 'meat processing'.

Owing to horrendous complexity and bureaucracy, only 17 categories remain. But researchers must still identify one category — and one only — for their grant application and outline the expected outcome of their research. This has encouraged them to pursue low-risk research rather than long-term fundamental research, Black said. It has also made it hard for young researchers to enter new fields. Indeed, lack of financial support for graduate students and postdoctoral fellows has meant that many young scientists have moved to Australia, where they are eligible for government grants.

Yet there is a fear that Australia may be headed down a similar path. Grant proposals and research missions in Australia now have to be justified in terms of socioeconomic benefits. "Demands for socioeconomic relevance lead to prostitution of the peer review process", warned Barry Osmond, director of ANU's Research School of Biological Sciences: "I am concerned that in our hasty opportunism, we all too willingly acquiesce in demands for assignments of priorities, adherence to strategic plans and so forth, knowing full well these are not the essence of progress in research and are inimical to creativity".

Gus Nossal, director of the Walter and Eliza Hall Institute of Medical Research in Melbourne, echoed this view. As a member of the council of Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO), he has repeatedly argued for the CSIRO to devote 25–30 per cent of its efforts to 'blue-sky' research. He claims "wide support" from the council, including members from industry. But the CSIRO seems to play down its contributions to blue-sky research when explaining its activities to the public and politicians. "Why be apologetic?", he appealed. "Let's argue very strongly that blue-sky research is very much in the national interest."

Peter Cook, Australia's Minister for Industry, Science and Technology, was (not surprisingly) more upbeat. He pointed out that in 1981 Australia ranked sixth (behind New Zealand) in public-sector OECD expenditure on research and development as a proportion of gross domestic product (GDP); now Australia is fourth, with New Zealand fourteenth: "Whether or not science policy is in the doldrums elsewhere, it has definitely been experiencing healthy winds of change and commitment by government in Australia". He singled out the Cooperative Research Centres (CRCs) set up between universities, CSIRO and industry in the early 1990s as an example of a "world-class innovation in science policy... linking public-sector research with business and other users of research". There are now more than 60 such centres.

But some Australian university researchers and officials are far from convinced about the benefits of the CRCs. They feel CRCs are a parasitic drain on university resources, despite government figures suggesting that universities gain financially from the centres, and have

refused to get involved with them.

Cook warned that scientists should not expect governments to fund all science: they must find "creative" solutions to attract private funds. Echoing this view, John Maddox, then editor of *Nature*, in a video link to the conference from London, pointed to China, where government institutes and universities have set up venture businesses to create new sources of revenue (see *Nature* 378, 537; 1995). But he warned that this model might not work in developed countries as many of the ventures in China are unlikely to last when the Chinese economy becomes more mature.

The situation in Japan stands in stark contrast with the rest of the Asia-Pacific region. Whereas governments in most of the region are aiming for strategic research, the Japanese government has just passed legislation to devote more resources to basic research, said Muneaki Samejima of Japan's opposition New Frontier Party.

Akito Arima, president of Japan's Institute of Physical and Chemical Research and a prime mover behind recent reforms of Japan's university system, explained that Japan today produces only about 500 PhDs a year. Furthermore, there is a greater emphasis on engineering than in the West.

Arima said that Japan is putting more effort into basic science: Japanese industry, he explained, is finding it increasingly difficult (and expensive) to gain access to intellectual property rights of other countries. And the development of new technology from basic research is now so fast that Japan must have its own indigenous basic research to stay competitive.

Perhaps the strongest argument for the support of basic research, however, came from Eugene Wong of Hong Kong University of Science and Technology. Wong showed that those countries investing heavily in basic research over many years are getting the best return on their total public reinvestment. Singapore reinvests nearly 40 per cent of its GDP, with only about 6 per cent average growth in GDP. Similarly, Japan reinvests 19 per cent with typically only 2–3 per cent growth. But the United States, with only 10 per cent reinvestment, has 2–3 per cent growth. Basic research, it seems, provides a base for development of new technologies and economic growth. □

*Copies of most of the conference speeches are available at the following website address: <http://biology.anu.edu.au/Pages/Pubs/NatConf/Nathome.html>

David Swinbanks is *Nature's* Japanese publisher and Asia-Pacific Editor.

New chapter for the fat controller

James Scott

The identification of a receptor in the brain of mice that binds the fat-regulating hormone leptin is a further step towards understanding — and perhaps treating — obesity.

LATE in 1994, this journal carried reports of experiments by Zhang *et al.*¹ that characterized the mouse *obese* gene; the product of this gene, dubbed leptin, is a hormone that is centrally involved in regulating fat tissue. Writing in *Cell*², Tartaglia and colleagues now describe the identification of a high-affinity receptor for leptin in the brain. Bit by bit, then, the process of leptin signalling and the control of obesity is being pieced together.

Obesity is the most prevalent nutritional disorder in consumer societies (for instance, one-third of North American adults are more than 20% overweight). The condition is associated with health problems such as non-insulin-dependent diabetes mellitus (NIDDM), hypertension, hyperlipidaemia, coronary heart disease and musculoskeletal disorders. Both environmental and genetic factors contribute, but in humans the genetic component is poorly defined.

Hence the importance of experiments with mice, in which there are several single-gene obese phenotypes. The *obese* (*ob*) mouse has a recessive mutation that results in severe obesity and NIDDM. The product of the gene, when not perturbed by mutation, is secreted from fat cells and acts on the hypothalamus of the brain to regulate the amount of body fat through the control of appetite and energy expenditure (see figure). Another, nearly identical obesity phenotype, called *diabetes* (*db*), affects the reception of the leptin signal in the brain. In overweight humans and rodents not defective for leptin, resistance to the leptin signal is more important than defective leptin production as a cause of excessive fat³⁻⁶. So identification of the hormone's target by Tartaglia *et al.* is a notable event.

The search for 'satiety factors' has a long history (reviewed in ref. 7). More than 40 years ago, Kennedy postulated the 'lipostat' theory. When an animal overeats, the extra fat somehow signals to the brain that the body is overweight, causing the animal to eat less and expend energy. The result is that the body fat can be regulated to within 1% over many years.

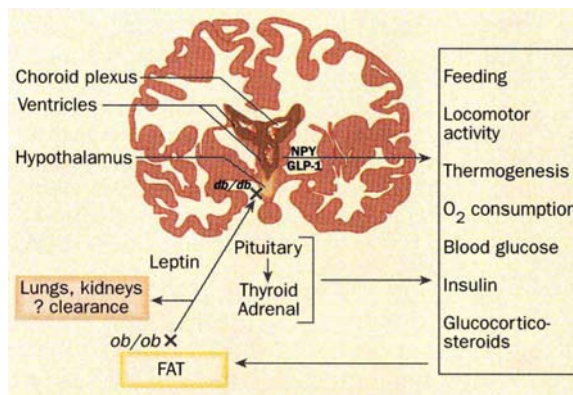
Strong support for the theory comes from various experiments. Damage to the hypothalamus produces a grossly overweight animal similar to genetically obese *ob/ob* or *db/db* mice (that is, with homozygous recessive mutations). Removal of fat leads to eating until fat stores are increased to compensate. Excessive feeding of one mouse in a parabiotic pair (in which the bloodstream is surgically joined) causes decreased feeding and weight loss by the underfed member of the pair. Damage to the hypothalamus in

is regulated by insulin and glucocorticoids. The hormone's site of action is the ventro-medial hypothalamus, where it decreases the biosynthesis and secretion of the neurotransmitter neuropeptide Y (NPY) — the most potent known stimulator of appetite.

In normal animals leptin suppresses food intake, and increases activity and heat production. It corrects the metabolic disturbances in the *ob/ob* mouse but has no effect on the *db/db* mouse, indicating that in this other recessive mutant there is a defect in the leptin receptor or in the post-receptor signalling pathway. In obese humans, levels of leptin in the blood plasma are normal or even high. Again, the implication is that it is not leptin production that is affected³⁻⁶; rather, human obesity appears to be caused by failure of leptin to reach its target in the brain, a consequence of defects in the receptor or in post-receptor processing.

So much for the background. Tartaglia and colleagues² used enzymatically and radiolabelled leptin to identify the hormone's binding sites in mouse choroid plexus (the site of cerebrospinal fluid production and transport of substances from the blood to the brain). Screening of a library of complementary DNA from the choroid plexus led to the identification of cDNAs encoding a leptin-binding protein (designated OB-R); this, it turns out, is a class 1 cytokine receptor that is most closely related to gp130, the signal-transduction component of the interleukin-6 receptor, and to the receptor for granulocyte colony-stimulating factor and leukaemia-inhibitory factor. The messenger RNA encoding this receptor is highly expressed in the choroid plexus, lungs and kidneys, and in the hypothalamus, leptin's site of action.

The OB-R gene itself lies quite close to the *db* gene, but binding of leptin in *db/db* mouse brain sections is almost normal and no mutation has been identified in the *db* mouse gene. A possible explanation suggested by this finding is that the cytoplasmic tail of the mouse leptin receptor is short compared to that of related receptors and of its human homologue. As Tartaglia *et al.* suggest, the OB-R gene may have a mutation that causes



Neurohormonal mechanisms of weight control. Leptin is secreted by fat cells in response to nutritional status, insulin and glucocorticoids. The leptin receptor (red) is found in the choroid plexus, hypothalamus, lungs and kidneys. The choroid plexus receptor presumably transports leptin to the brain. The hypothalamic receptor receives the leptin signal and controls energy expenditure through the balance of activity between appetite suppressing (GLP-1) and stimulatory (NPY) neurotransmitters. Leptin receptors in the lung and kidney probably dispose of excess leptin. Mouse leptin (*ob*) and leptin receptor (*db*) mutations are shown.

one of a parabiotic pair results in excessive weight gain in the affected mouse, with reduced feeding and loss of weight in the other. An *ob/ob* mouse as one of a parabiotic pair with a normal animal eats less and gains less weight, whereas a normal mouse paired with a *db/db* mouse shuns food and starves to death. Finally, an *ob/ob* mouse joined to a *db/db* mouse reduces its food intake and loses weight. All of these results are consistent with the production of a circulating hormone that, like insulin, is effective over long periods.

Leptin, then, emerged as the long-sought 'satiety hormone'⁸⁻¹². It is secreted exclusively from fat cells in response to an increase in fat mass and acts over several hours; its production is suppressed by feeding and enhanced by starvation; and it

abnormal splicing of the intracytoplasmic tail of the receptor.

How does leptin enter the brain, given that the blood-brain barrier is normally impervious to polypeptide hormones? The simple answer may be that, like insulin, it is transported into the brain by the choroid plexus; the intense expression of OB-R there suggests that is indeed the case. The high level of OB-R expression in the lungs and kidneys, sites of polypeptide hormone disposal, points to these tissues as the places where excess leptin is dealt with.

The past year has provided other insights into the control of body weight. Prostaglandins of the J series have been identified as endogenous activators of peroxisome proliferator-activated receptors, so-called orphan nuclear receptors that are implicated in fat-cell homeostasis^{13,14}. The antidiabetic thiazolidinedione drugs have been found to act on gene transcription through this same nuclear-receptor signalling pathway, so providing a route for the control of metabolism in fat cells. Finally, in a paper published last week¹⁵, Turton and colleagues have shown that injection into the brain of the neurotransmitter glucagon-like peptide 1 (GLP-1) inhibits feeding in rats. The activity of GLP-1 can be blocked at its receptor by a natural antagonist called exendin, which augments the response to the appetite stimulant NPY. GLP-1 therefore appears to be a central mediator of satiety.

Much of course remains to be done. How is the biosynthesis and secretion of leptin controlled? Is the newly described protein indeed the true leptin receptor? How is the leptin signal transmitted to NPY and beyond? Does overeating with time disrupt leptin signalling? What is the precise role of GLP-1 and its receptor in satiety? And are other neurotransmitters involved?

These issues are no doubt being pursued with great vigour. Not least, at stake is a potential billion-dollar-a-year prize for molecular-level treatments for obesity. Such therapies might take the form of drugs that enhance leptin production or otherwise amplify the leptin signal. □

James Scott is in the Department of Medicine, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK.

1. Zhang, Y. *et al.* *Nature* **372**, 425–432 (1994).
2. Tartaglia, L. A. *et al.* *Cell* **83**, 1263–1272 (1995).
3. Considine, R. V. *et al.* *J. clin. Invest.* **95**, 2986–2988 (1995).
4. Hamilton, B. S. *et al.* *Nature Med.* **1**, 953–956 (1995).
5. Lonnqvist, F. *et al.* *Nature Med.* **1**, 950–953 (1995).
6. Maffei, M. *et al.* *Nature Med.* **1**, 1155–1161 (1995).
7. Rink, T. J. *Nature* **372**, 406–407 (1994).
8. Campfield, L. A. *et al.* *Science* **269**, 546–549 (1995).
9. Halaas, J. L. *et al.* *Science* **269**, 543–546 (1995).
10. Pelleymounter, M. A. *et al.* *Science* **269**, 540–543 (1995).
11. Saladin, R. *et al.* *Nature* **377**, 527–529 (1995).
12. Stephens, T. W. *et al.* *Nature* **377**, 530–532 (1995).
13. Forman, B. M. *et al.* *Cell* **83**, 803–812 (1995).
14. Rameh, L. E. *et al.* *Cell* **83**, 821–830 (1995).
15. Turton, M. D. *et al.* *Nature* **379**, 69–72 (1996).

Ediacaran in deep water

Douglas Palmer

ALMOST 50 years ago the first description of mysterious Precambrian fossils from the Ediacara Hills in Australia was published¹. Subsequent discoveries from elsewhere in the world led to the definition of the so-called Ediacaran fauna of jellyfish-like organisms. These were globally distributed and apparently became extinct well before the beginning of the Cambrian period, here taken to be about 540 million years (Myr) ago, and the evolutionary 'explosion' of multicellular animals with shells. The fossils have been the subject of much discussion as regards their taxonomic affinities and evolutionary history.

Debate enters a fresh phase with the description by Crimes *et al.*² of new examples of Ediacaran fossils from deep-water deposits, some 510 Myr old³, in County Wexford, Ireland. Taken together with a paper by Grotzinger and colleagues⁴, which provided dates for Namibian specimens from close to the Precambrian–Cambrian boundary, the finding may well be the death knell for the idea that the Ediacaran fauna was a failed evolutionary experiment. Moreover, the unusual preservation of the Irish fossils provides support for the view⁵ that, in life, the creatures were not soft-bodied but had a rigid outer wall unlike any extant group.

The Ediacara are dominated by a few forms, which look superficially like living jellyfish and seapens of the phylum Cnidaria. But these 'jellyfish' seem to lack certain anatomical details characteristic of living species, such as any clear indication of a mouth or gastric cavity, and their three-dimensional preservation as sediment-filled moulds has raised serious questions about their taxonomic affinities. The taphonomy, or mode of preservation, of such 'soft-bodied' organisms is a key issue. Most of the Ediacaran fossils have until now been found in sandstones originally laid down in shallow water. Somehow or other, the three-dimensional form of the medusoid-like fossils is generally preserved without substantial flattening but with the body shape infilled with sediment. No trace of tissue, mineralized or not, remains. The question is whether such preservation can be produced from tissues like those found in living medusoids, or whether it indicates that we are dealing with some quite different type of organism.

A feature of the Irish fossils is that they occur in turbiditic deposits laid down in deep water. There are two different medusoid-like taxa, both preserved as sediment-filled moulds. The small (2–8 mm diameter) discoid *Nimbia occlusa* fossils occur with similar orientation flat on a single bedding plane. The larger *Ediacara booleyi* discs are more variable in size

(26–200 mm in diameter) and occur jumbled together at all angles within several successive strata.

The unfragmented preservation of the specimens shows that their body tissue was tough enough to resist being torn apart while being tumbled about in fast-moving, bottom-hugging, turbulent density currents — the submarine equivalent of a washing machine full of sand. Furthermore, entombment in density currents may be telling us something significant about the organisms' own density. Would modern jellyfish be buried in turbidite deposits like these? I doubt it, but there is clearly a need for more data on this score.

What few taphonomic experiments⁶ there have been on comparable soft-bodied living cnidarians show that, on decay, jellyfish tend to leave flat impressions. Indeed, it has been argued⁷ that the typical Ediacaran sandstone preservation is more akin to that found in plant fossils with cellulose-strengthened stem tissues. The Irish fossils, with their different sediment matrix, show that three-dimensional preservation of these organisms was not specific to shallow-water sandstones.

On the palaeobiology of the new specimens, Crimes *et al.* argue that the absence of any indication of a mouth is good evidence against a cnidarian affinity. But to my mind the jury must remain out on that question, especially in view of decay experiments⁸ on cnidaria which showed that such oral structures are not necessarily preserved in medusoids. Finally, it may be that the habitat of these Ediacaran 'medusoids' was the result of one of the earliest examples of competitive exclusion — as the authors point out, it looks as if they retreated from their original shallow-water niche into a deep-water refuge.

We are no doubt still seeing only the opening shots in the revived Ediacaran debate. On the taxonomic front, more information should come from detailed comparative study of the diagnostic features of the new fossils and other Ediacaran species; and we can hope for further, illuminating discoveries, especially in the beautifully preserved, flat-lying Late Precambrian strata of Namibia. □

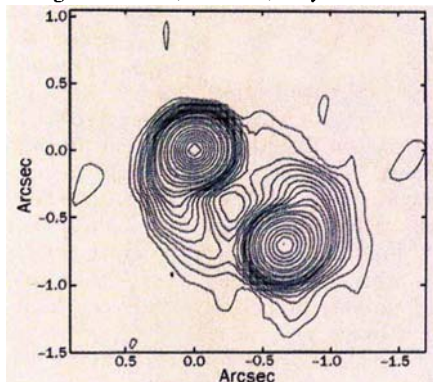
Douglas Palmer is at 31 Mawson Road, Cambridge CB1 2DZ, UK.

1. Sprigg, R. C. *Trans. R. Soc. S. Aust.* **71**, 212–224 (1947).
2. Crimes, T. P., Inssole, A. & Williams, B. P. *J. Geol.* **30**, 89–109 (1995).
3. Moczyłowska, M. & Crimes, T. P. *Geol. J.* **30**, 111–128 (1995).
4. Grotzinger, J. P., Bowring, S. A., Saylor, B. Z. & Kaufman, A. J. *Science* **270**, 598–604 (1995).
5. Seilacher, A. *J. Geol. Soc. Lond.* **149**, 607–613 (1992).
6. Briggs, D. E. G. *Palaio* (in the press).
7. Retallack, G. J. *Paleobiology* **20**, 523–544 (1994).
8. Norris, R. D. *Lethaia* **22**, 381–393 (1989).

The optics of cosmology

C. S. Kochanek

GRAVITATIONAL lenses promise to answer many of the most important questions in cosmology and in galactic structure and evolution. Or so astronomers have claimed for the past fifteen years. There has been progress. Lenses provide the strongest constraints on the cosmological constant (a measure of the self-repulsion of space), agree with other estimates of the Hubble constant and strongly suggest the existence of dark matter in the inner regions of elliptical galaxies. Yet, all in all, they have failed



An 8.4-GHz VLA map³ of the lensed radio source PKS1830-211, showing the two main components of the ring-like image.

to fulfil their early promise. It is not, I think, a failure in our lenses, but in the psychology of the field. To paraphrase one prominent astronomer, most observers are more interested in the numerator (finding new lenses) than in the denominator (accurate, quantitative studies of existing lenses and their source populations). Fortunately, not all observers have adopted this view; a good example is the clever approach of Wiklind and Combes — reported on page 139 of this issue¹ — who determine the redshift of the lens responsible for the PKS1830-211 radio ring^{2,3} (see figure) by searching for molecular absorption lines.

The gravity of an intervening galaxy can magnify and produce multiple images of distant objects. Twenty-three convincing examples of these lenses have been found, where the sources are quasars, radio sources or distant galaxies⁴. Only six have measurements of both source and lens redshift, which are needed to determine the distances to the lens and the source, as well as physical parameters such as galaxy masses and the Hubble constant, and to test cosmological models. Two promising new approaches to finding them are infrared⁵ and radio spectroscopy. Although the lens redshift of B0218+357 was confirmed by detection of 21-cm absorption⁶, that of PKS1830-211 is the first to be determined by radio spectroscopy.

Detailed photometry of lensed systems is largely confined to projects that monitor

the variability of lensed images. In particular, source fluctuations appear in the images at different times due to differences in path length, and these delays can be used to determine the Hubble constant geometrically, with few of the interminable calibration problems of local estimates. A delay is measured^{7,8} in the first lens to be discovered, 0957+561, but this system is a compound lens consisting of a galaxy and a cluster, complicating the interpretation. Most lenses avoid the difficulties of 0957+561, but it has been impossible to obtain the necessary radio or optical telescope time to systematically monitor the better candidates. PKS1830-211 and B0218+357 show strong radio variability, and several quasar lenses show low-level optical variability. A small fraction of the observational resources devoted to traditional means of measuring the Hubble constant would allow this field to explode.

The precise redshift of $z=0.886$ measured by Wiklind and Combes might allow us to determine the Hubble constant at an intermediate cosmological scale. They used a 15-m radio telescope to look for absorption in the 2- and 3-mm wavebands caused by the rotational lines of a few simple molecules such as CS and HCN. Because the line of sight to the source passes within a few thousand parsecs of the centre of the inferred lensing galaxy, there was some chance of finding a molecular cloud there. If, as the authors suggest, only one of the two main images is covered by a cloud, their signals can be distinguished by the presence of absorption lines, and it is possible to determine the relative fluxes of the two images using a single-dish telescope instead of an interferometer. In this way one could measure the time delay and, if the source redshift can be measured, the Hubble constant.

Good photometry of the lensing galaxies is less common, even though it can be used to test cosmological models (for example, how much matter is in the Universe, and whether there is a cosmological constant). We measure the apparent magnitude of the lens galaxy and the separation of the lensed

images, which is related to the mass of the lens and therefore to its absolute luminosity. For a given absolute luminosity, the apparent magnitude depends on the cosmological model. At a redshift of 0.8, a lens galaxy in a Universe dominated by self-gravitation of space (with the cosmological constant $\Lambda_0=1$) is 2.0 magnitudes fainter than in a flat, matter dominated model (with matter density $\Omega_0=1$). Corrections must be made for evolution, and the relation between luminosity and image separation must be calibrated, but these uncertainties are smaller than the cosmological differences. The dominant uncertainty is the extraordinarily mundane problem that most measurements of lens galaxy magnitudes do not report on the aperture through which the flux was measured! This single missing number can lead to uncertainties larger than the effects of cosmology.

In theory, the strongest cosmological constraints from gravitational lenses should come from surveys for radio lenses — these surveys have found many more lensed systems⁹⁻¹¹, and are immune to some systematic problems in the quasar surveys, such as dust. In practice, there is a crippling uncertainty about the redshift distribution of faint (< 100 mJy) radio sources. Radio source redshifts were popular until 1990, but interest waned rapidly as cosmology became the hot topic of the day. Yet the mean redshift of 50-mJy sources must be about 0.5 if the cosmological constant is large, and about 1.5 if it is small — a simple and clean way of attacking a fundamental cosmological problem through the neglected 'denominator'.

The promises made fifteen years ago still hold — gravitational lenses are powerful constraints on cosmology and galactic structure. The fact remains, however, that a fundamental bottleneck in using gravitational lenses as tools in astrophysics is the focus on finding new ones, rather than making precise measurements on the known lenses. Any progress, such as the PKS1830-211 lens redshift, is a step in the right direction, and with enough such steps we can fulfil the earlier promises. □

C. S. Kochanek is at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA.

- Wiklind, T. & Combes, F. *Nature* **379**, 139-141 (1996).
- Rao, A. P. & Subrahmanyam, R. *Mon. Not. R. astr. Soc.* **231**, 229-236 (1988).
- Jauncey, D. L. et al. *Nature* **352**, 132-134 (1991).
- Keeton, C. R. & Kochanek, C. S. in *Astrophysical Applications of Gravitational Lensing*, Proc. 173 Symp. of the IAU (eds Kochanek, C. S. & Hewitt, J. N.) 419-439 (Kluwer, Dordrecht, in the press).*
- Lawrence, C. R. in *Astrophysical Applications of Gravitational Lensing*, Proc. 173 Symp. of the IAU (eds Kochanek, C. S. & Hewitt, J. N.) 299-304 (Kluwer, Dordrecht, in the press).*
- Carilli, C. L., Rupen, M. P. & Yanny, B. *Astrophys. J.* **412**, L59-L62 (1993).
- Haarsma, D. B., Hewitt, J. N., Burke, B. F. & Lehar, J. in *Astrophysical Applications of Gravitational Lensing*, Proc. 173 Symp. of the IAU (eds Kochanek, C. S. & Hewitt, J. N.) 43-44 (Kluwer, Dordrecht, in the press).*
- Schild, R. E. & Thomson, D. J. in *Astrophysical Applications of Gravitational Lensing*, Proc. 173 Symp. of the IAU (eds Kochanek, C. S. & Hewitt, J. N.) 51-52 (Kluwer, Dordrecht, in the press).*
- Burke, B. F., Lehar, J. & Connor, S. R. in *Gravitational Lenses* (eds Kayser, R., Schramm, T. & Nisner, L.) 237-244 (Springer, Berlin, 1992).
- Patnaik, A. R. in *Gravitational Lenses in the Universe* (eds Surdej, J., Fraipont-Caro, D., Gosset, E., Refsdal, S. & Remy M.) 311-321 (Univ. Liège, 1994).
- Meyers, S. T. in *Astrophysical Applications of Gravitational Lensing*, Proc. 173 Symp. of the IAU (eds Kochanek, C. S. & Hewitt, J. N.) 317-322 (Kluwer, Dordrecht, in the press).*

*Conference proceedings are available at <http://cfata2.harvard.edu/iau173/book.html>

Pinpointing the centre

Mary F. Lyon

THE mechanism of inactivation of one of the two X chromosomes in somatic cells of female mammals has remained an enigma since its discovery in 1961. This inactivation is necessary because, unlike males, who carry one X and one Y chromosome, females carry a double dose of X-linked genes. Now, in a breakthrough announced on page 131 of this issue, Penny *et al.*¹ pro-

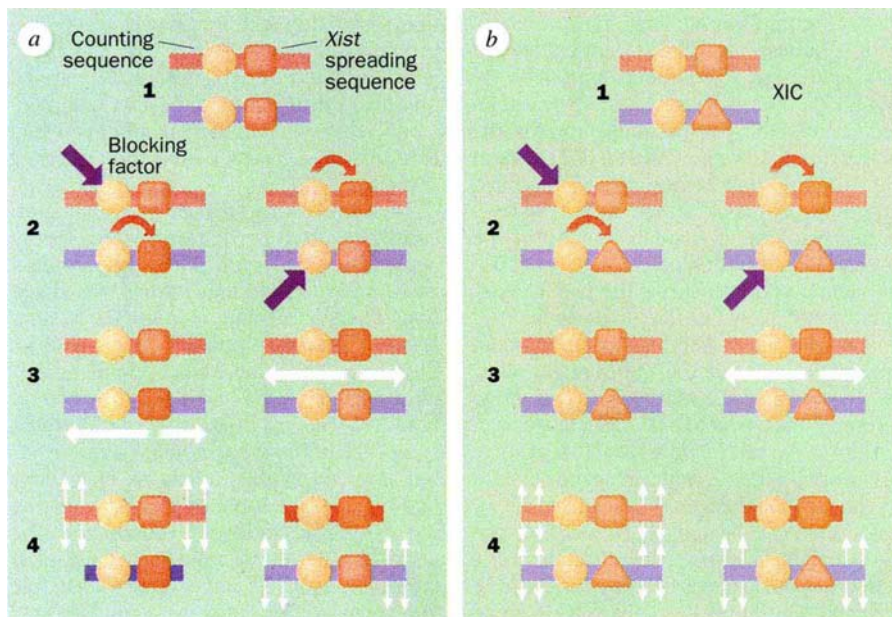
The concept of an X-inactivation centre (XIC) arose from the finding that when the X chromosome was broken into two parts and rejoined with parts from another, non-sex, chromosome (autosome), only one of the two broken X-chromosome segments underwent inactivation. The interpretation was that the inactivation was initiated from a special site, the X-inactivation centre, on

nism that maintains one X active per two autosome sets and which is postulated to act by blocking one XIC. Hence, the XIC is thought to take part in counting and in the choice of X chromosome for activity, and in the promulgation of the signal for spreading along the chromosome (see figure).

The first breakthrough came with the discovery in humans of the *XIST* gene⁴, which mapped to the XIC and had the unique property of being expressed by the inactive but not the active X chromosome. The corresponding gene *Xist* was quickly cloned in the mouse^{5,6}. It is expressed in the early embryo well before the time of initiation of inactivation⁷, and hence this expression could be causal. Furthermore, its expression in embryonic stem (ES) cells is appropriate. When ES cells that are chromosomally XX are maintained in the undifferentiated state, both X chromosomes remain active and *Xist* is expressed at a very low level; however, when they are allowed to differentiate, X inactivation occurs and *Xist* expression increases markedly⁷. Thus, from the combination of its location and its expression in adults, embryos and ES cells, *XIST/Xist* has the appropriate properties to fulfil the functions of the XIC.

Penny *et al.*¹ have now provided clear evidence that transcription of the *Xist* gene is indeed necessary for the inactivation of the X chromosome that bears it. They knocked out seven kilobases of the DNA in the first coding region (exon) of *Xist* in ES cells and showed that this destroyed its activity (created a null allele). The ES cells were heterozygous for X chromosomes with distinguishable alleles of *Xist* and other X-linked genes, and thus it was clear which X carried the knockout. When the cells were allowed to differentiate, X inactivation occurred, as manifested by an asynchronously replicating chromosome, but only the X not bearing the knockout underwent inactivation, as shown by the expression of the X-linked genes *Pgk1*, *Rps4x* and *SmcX*.

Thus, the counting mechanism still recognized the XIC with the null *Xist* allele and the authors' suggested interpretation is that either the normal or the knockout X could be selected to remain active, with a probability depending on the different *Xce* alleles they carried. If the knockout X was selected to remain active, the other X underwent inactivation normally. However, if the normal X was selected, then the knockout X failed to become inactive and the cell then had two active X chromosomes (see figure).



Hypothetical sequence of events at initiation of X inactivation in normal cells (a) and those with deletion of *Xist* (b). The XIC includes sequences involved in counting and choice of X for inactivation (circle) and in spreading of inactivation (square). One counting sequence interacts with an unknown *trans*-acting blocking factor (large arrow in a2), the spreading sequence (*Xist*) in the other XIC is activated (red arrow) and *Xist* is transcribed. Inactivation then spreads in both directions from the unblocked XIC (a3) and this X chromosome becomes condensed and inactive (a4), while the X with the blocked centre is transcribed normally (a4). In cells with a deletion of *Xist* (triangle), the *trans*-acting factor interacts normally with the counting sequence (b2), but the deleted *Xist* is not transcribed. In cells with the deleted *Xist* gene blocked (b2, right), the normal XIC initiates inactivation (b3, right), but in cells with the normal *Xist* gene blocked (b2, left) the deleted *Xist* is not transcribed and both X chromosomes remain active (b4, left). *In vivo*, cells of this type would be rapidly eliminated by cell selection. Another possible interpretation of Penny *et al.*'s data¹ is that in all cells with the deleted *Xist* the counting mechanisms block the deleted XIC only.

vide crucial evidence that the *Xist* gene (for X-inactive specific transcript) forms an essential part of the X-inactivation centre, the source of the inactivation which spreads along the chromosome.

During X inactivation (reviewed in ref. 2), the two X chromosomes in any cell of female mammals become differentiated early in development. One, the inactive X chromosome, assumes a compact, condensed structure, replicates its DNA asynchronously, and is not transcribed into messenger RNA. The other, active, X chromosome has the more open structure characteristic of functional chromatin and is transcribed normally. This is a form of dosage compensation, ensuring equal gene dosage of X-linked genes in males and females.

the X chromosome, spreading from there in both directions (see figure), and that segments without this centre could not undergo inactivation³.

By studying different translocations and X-chromosome deletions, the XIC was mapped to a small region on both mouse and human X chromosomes. In addition, in the mouse a factor termed *Xce* (X-controlling element), which affects the choice of X for inactivation, also maps to the region of the XIC. In animals homozygous for any allele of *Xce*, each X chromosome is inactivated in a 50:50 ratio, whereas in those heterozygous for different alleles the ratio might be 60:40 or 70:30 depending on the 'strength' of the two alleles.

The number of active X chromosomes in any cell is determined by a counting mecha-

1. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S. & Brockdorff, N. *Nature* **379**, 131–137 (1996).
2. Lyon, M.F. *Adv. Genome Biol.* **4**, 119–151 (1995).
3. Rastan, S. *J. Embryol. exp. Morph.* **78**, 1–22 (1983).
4. Brown, C.J. *et al.* *Nature* **349**, 38–44 (1991).
5. Borsani, G. *et al.* *Nature* **351**, 325–329 (1991).
6. Brockdorff, N. *et al.* *Nature* **351**, 329–331 (1991).
7. Kay, G.F. *et al.* *Cell* **72**, 171–182 (1993).
8. Brown, C.J. & Willard, H.F. *Nature* **368**, 154–156 (1994).
9. Brockdorff, N. *et al.* *Cell* **71**, 515–526 (1992).
10. Brown, C.J. *et al.* *Cell* **71**, 527–542 (1992).

The authors also studied the effect of the *Xist* knockout *in vivo* in chimaeric embryos, made by aggregating ES cells carrying the knockout with normal 8-cell embryos. Again, the counting mechanism operated and X-inactivation occurred, and once more only the X with the normal *Xist* allele underwent inactivation. In contrast to the ES cells, there was no evidence of cells with both X chromosomes active — in line with earlier work showing that cells with excess X-chromosome activity *in vivo* are rapidly eliminated by cell selection. Thus, this work provides clear evidence that transcription of *Xist* is required for the spreading of inactivation along the X chromosome carrying it, and one of the burning questions, the identity of the XIC, is solved.

An exciting host of new questions is now opened up for study. How does the counting mechanism work? Suggestions range from a limited amount of a specific molecule to the location of the active X at a single site in the nucleus. The fact that counting still occurred despite the knockout could mean that *Xist* is not needed for counting or that other regions of the gene are involved. It is interesting that the counting and spreading functions of the XIC have been to some extent separated, because counting may be a later evolutionary devel-

opment found only in eutherians. In marsupials, the X chromosome for inactivation is chosen by parental origin and counting may not occur. Similarly, in X inactivation in male germ cells no counting is involved. Does *Xist* function in marsupials and male germ cells? What is the significance of the expression of *Xist* in the adult? In cells that have already undergone X-inactivation, loss of the *Xist* gene does not cause reactivation⁸. Thus, the significance of expression of *Xist* in adult tissues is not clear.

Even more fascinating is the question of how spreading is brought about. It is a long-range process apparently operative over many megabases, but with some form of local response in that some genes escape inactivation. The *XIST/Xist* gene does not code for a protein and its 15–17-kilobase transcript is retained in the nucleus and is associated with the inactive X chromosome itself^{9,10}. How does it function? Further and more subtle knockouts of *Xist* will no doubt answer some of the questions, but a few more advances will be needed before the mechanism of X inactivation is thoroughly understood. □

Mary F. Lyon is in the MRC Mammalian Genetics Unit, Harwell, Didcot, Oxfordshire OX11 0RD, UK.

CRUSTAL RECYCLING

Best friend hides deep secret

William M. White

THE commercial value of diamonds increases with their purity and clarity. Their scientific value, however, often lies in the mineral inclusions that can occur within them. On page 153 of this issue, Daniels and others¹ report finding an inclusion of staurolite, a mineral typically found in metamorphosed clay-rich sediments. They argue that because diamond

only forms at pressures reached 120 km or more into the Earth, this is tangible confirmation of what geologists have suspected from isotopic studies for some time — that material from the continental crust is recycled into the mantle.

Once encapsulated in diamond, a mineral is isolated from external influences. Inclusions in diamonds carried to the surface by volcanic eruptions

thus provide geologists with a rare glimpse into the mantle. The discovery by Daniels *et al.* is only the most recent of a series of similar ones made over the past 10 to 15 years. Richardson and colleagues² found inclusions as old as 3.2 billion years, demonstrating not only the great antiquity of some diamonds, but also the great stability of some parts of the mantle underlying the continents. Indeed it appears that once a continental mass forms, a region of mantle down to depths of 200 km or more (the subcontinental lithosphere) is stabilized, or 'frozen in', below it. Other diamonds contain inclusions

that appear to come from the lower mantle, the region deeper than 650 km^{3,4}. These may be the only samples we have of the Earth's deep interior.

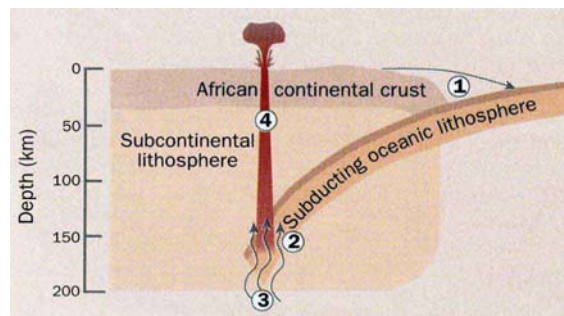
Staurolite is a common mineral in aluminium-rich sediments metamorphosed at moderate temperatures and pressures, and has never before been reported in mantle rocks. The staurolite inclusion found by Daniels and co-workers is compositionally very similar to those found in metamorphosed sediments. Thus it seems most likely that this staurolite crystal formed within metamorphosed crustal rocks which were subsequently carried into the mantle through subduction.

In higher-pressure metamorphic rocks, staurolite is usually replaced by a combination of kyanite and garnet, raising the question of how this particular staurolite could survive to pressures where it could be encapsulated by diamond. Daniels *et al.* speculate that it may have formed within rocks lacking the free SiO₂ (quartz or its high-pressure equivalents) that is required for these reactions. Alternatively, they suggest it may have been included within another mineral, such as garnet, protecting it from reaction until it could be encapsulated by diamond.

It has been known for some time that the carbon isotope composition of terrestrial diamonds is bimodal⁵. Most diamonds are about 5‰ poorer in ¹³C than typical carbonate rocks (carbon isotope compositions are usually reported as parts per thousand (‰) deviations from Pee Dee belemnite, a calcium carbonate standard; these values are denoted as δ¹³C_{PDB}). This δ¹³C_{PDB} value of −5‰ is the same as that found in basalts and carbonatite magmas and probably represents the isotopic composition of mantle carbon. A smaller fraction of diamonds have lower δ¹³C_{PDB}, ranging from −10 to −35‰, with the most common value being around −20‰. The isotopic composition of low-δ¹³C_{PDB} diamonds is similar to that of organic carbon at the Earth's surface, giving rise to speculation that these diamonds consist of carbon that was once there. Interestingly, however, the diamond studied by Daniels *et al.* seems to have formed from mantle carbon, judging by its δ¹³C_{PDB} value of −5‰.

The low-δ¹³C diamonds often contain inclusions of minerals commonly found in eclogite, a rock produced when basalt, a common igneous rock, is subjected to high pressure. The oceanic crust consists primarily of basalt and would transform to eclogite when it is subducted into the mantle. This is viewed by some as further evidence that low-δ¹³C diamonds represent material recycled from the Earth's surface. Other evidence includes the isotopic and chemical composition of sulphide inclusions in diamonds and the isotopic composition of nitrogen in low-δ¹³C diamonds.

Other minerals typically found in meta-



A possible route for the formation of staurolite inclusions in diamond. (1) Clay-rich sediments from the continents are carried into the mantle in an ancient subduction zone. Staurolite forms as temperature and pressure increase. (2) Subduction ceases and the subducted oceanic lithosphere becomes frozen into the newly developed subcontinental lithosphere beneath Africa. (3) Invasion of a carbon-bearing fluid. Diamond nucleates on a staurolite crystal, eventually encapsulating it. (4) Diamond is carried to the surface in the violent eruption of a kimberlite magma.

morphic rocks of the Earth's crust, including kyanite, coesite, corundum and alkali feldspars^{6,7}, have also been reported as inclusions in diamonds. These might also have formed in the crust or from crustal precursors, though that is less certain than for staurolite. At one point, the oxygen isotope compositions of lavas from the Pitcairn seamounts appeared to provide convincing evidence for the presence of recycled crustal material in mantle plumes⁸, but the failure by other laboratories to reproduce the data has placed those results in serious question⁹.

Thus, as Daniels and colleagues point out, there is already evidence that crustal material can be carried deep into the Earth's mantle. Yet much of this evidence remains controversial. The discovery of a staurolite crystal in diamond is perhaps the most tangible and unambiguous evidence to date in this respect.

Subduction, the process by which the lithospheric plates formed at mid-ocean ridges are returned to the mantle, is an integral part of plate tectonics and the Earth's behaviour. If sediment is subducted as well, this process can return continental crust to the mantle, as Armstrong¹⁰ pointed out early on. So finding material of crustal provenance in the mantle should not be surprising. What remains uncertain is the geological significance of this process. Has this process balanced crustal production and so maintained a constant continental mass, as Armstrong argued? Is it, as others have argued, the process that produces much of the chemical heterogeneity observed in the mantle? It now appears that the subcontinental lithosphere is particularly heterogeneous, perhaps no less so than the crust above it. To what degree are subduction and the processes associated with it responsible for this variability? Future studies of diamonds, those almost indestructible capsules carrying mineralogical and chemical clues from the Earth's deep interior, may help provide the answers. □

William M. White is in the Department of Geological Sciences, Cornell University, Ithaca, New York 14853, USA.

Spiralling to destruction at the edge of chaos

Alain Karma

MANY pattern-forming systems in nature exhibit a transition from an ordered state to a chaotic 'turbulent' state. There are hardly any systems, however, where this transition has been understood in its fine details and where theoretical predictions have been quantitatively tested by experiment. On page 143 of this issue, Ouyang and Flesselles¹ demonstrate that in a special parameter range of the Belousov-Zhabotinsky (BZ) reaction, the transition from clock-like spiral-wave rotation to spiral-wave turbulence occurs exactly according to theoretical expectation.

Nonlinear dynamics has been quite successful in explaining the transition from periodic to aperiodic motion in systems with a few degrees of freedom. Period doubling and intermittency are well-known routes to chaos that have been seen in numerous experiments and that are now standard textbook material. In non-equilibrium systems that naturally form patterns, however, the transition from order to chaos involves the creation of disorder in both space and time, with many degrees of freedom that increase in number with the size of the system. This transition has remained difficult to describe theoretically, except in a few model examples.

One of these examples is the complex Ginzburg-Landau equation (CGLE). Despite its simplicity, this equation already exhibits an extremely rich spectrum of dynamical behaviour and for this reason has been an important focus for theoretical studies of spatiotemporal chaos². In the context of reaction-diffusion systems, it can quantitatively describe oscillatory media close to a Hopf bifurcation (the onset of an instability that oscillates in time). Wave patterns form easily in such media because different regions do not necessarily oscillate in phase with each other, and they are diffusively coupled (the pattern can travel between them). In a spiral wave the phase varies smoothly in space except at the spiral centre, where the contours of constant phase meet at a singular point, or topological defect, and the oscillation amplitude vanishes.

Oscillatory media described by the CGLE can exhibit an ordered state, with a single, stable spiral wave rotating at a fixed angular frequency³, or a disordered, turbulent state^{4,5}, with a seemingly random creation and annihilation of defect pairs (each defect being the centre of a small spiral) that have a well characterized average density⁵. The transition

between these two states occurs when the rotation frequency of the spiral is such that the waves propagating outwards from the centre become subject to an instability that causes a long-wavelength modulation. When the amplitude of this modulation becomes sufficiently large, waves break up and defect pairs are generated, leading the system into a turbulent state. Interestingly, this instability does not destroy the initial spiral pattern all at once, because of its 'convective' nature⁶; small perturbations propagate away from the spiral centre faster than they can grow in amplitude, so the central portion of the spiral remains stable. As a control parameter is changed, the growth rate of perturbations increases and this stable region shrinks. Eventually an absolute stability limit is reached where perturbations spread faster than they can propagate away from the centre. This theoretical picture is precisely the one that Ouyang and Flesselles have been able to reproduce experimentally in the BZ reaction, with many details matching theory convincingly.

Spiral waves are not new to the BZ reaction. They have been experimentally observed and modelled for decades, in parameter ranges where the spirals are usually stable² and hence not susceptible to breakup and turbulence (with perhaps some exceptions⁷). Ouyang and Flesselles, however, have studied the BZ reaction for high concentrations of sulphuric acid and in an open system where the reaction can be accurately maintained near a Hopf bifurcation for a long time, which is precisely the regime described by the CGLE. This is what has allowed them to validate experimentally a theoretical model for the transition to spatiotemporal chaos near a Hopf bifurcation. Their results reinforce our belief that amplitude equations such as the CGLE provide an accurate description of

1. Daniels, L. R. M., Gurney, J. J. & Harte, B. *Nature* **379**, 153–156 (1995).
2. Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198–202 (1984).
3. Scott-Smith, B. H., Danchin, R. V., Harris, J. W. & Stracke, K. J. in *Kimberlites and Related Rocks* (ed. Kornprobst, J.) 121–142 (Elsevier, Amsterdam, 1984).
4. Harte, B. & Harris, J. W. *Mineralog. Mag.* **58A**, 384–385 (1994).
5. Mathey, D. P. *Terra Cognita* **7**, 31–38 (1987).
6. Sobolev, N. V. *Zapiski vsesoyuznogo mineralogicheskogo Obshchestva* **112**, 389–397 (1983).
7. Otter, M. L. & Gurney, J. J. in *Kimberlites and Related Rocks Volume 2: Their Mantle/Crust Setting, Diamonds and Diamond Exploration* (eds Ross, J. et al.) 1042–1053 (Blackwell Scientific, Victoria, 1989).
8. Woodhead, J. D., Greenwood, P., Harmon, R. S. & Stoffers, P. *Nature* **362**, 809–813 (1993).
9. Eiler, J. M. et al. *Nature* **377**, 138–141 (1995).
10. Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).

1. Ouyang, Q. & Flesselles, J.-M. *Nature* **379**, 143–146 (1996).
2. Cross, M. C. & Hohenberg, P. C. *Rev. Mod. Phys.* **65**, 851–1112 (1993).
3. Hagan, P. S. *SIAM J. appl. Math.* **42**, 762–786 (1982).
4. Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).
5. Coulet, P., Gil, L. & Lega, J. *Phys. Rev. Lett.* **62**, 1619–1622 (1989).
6. Aranson, I. S., Aranson, L., Kramer, L. & Weber, A. *Phys. Rev. E* **46**, R2992–R2995 (1992).
7. Markus, M., Kloss, G. & Kusch, I. *Nature* **371**, 402–404 (1994).
8. Morris, S. W., Bodenschatz, E., Cannell, D. S. & Ahlers, G. *Phys. Rev. Lett.* **71**, 2026–2029 (1993).
9. Assenheimer, M. & Steinberg, V. *Nature* **367**, 345–347 (1994).
10. Jakubith S., Rotermund, H. H., Engel, W., von Oertzen, A. & Ertl, G. *Phys. Rev. Lett.* **65**, 3013–3016 (1990).
11. Winfree, A. T. *Science* **266**, 1003–1006 (1994).

spatially extended non-equilibrium systems near bifurcation points, where dynamical behaviour is expected to be universal.

Nature, however, rarely operates near such points unless artificially encouraged. In other systems showing defect-mediated turbulence, from fluid convection^{8,9} to surface catalysis¹⁰ to heart muscle¹¹, the transition from order to chaos occurs via what seem to be different wave instabilities or entirely different dynamical mechanisms. In most of these systems, and

particularly in heart tissue, where electrical-wave turbulence directly affects human health, theory and experiment are yet to meet in a clean and convincing way. A major challenge also lies ahead in characterizing the complex dynamical state that forms after the onset of turbulence. That requires the development of theoretical tools to describe the collective behaviour of defects. Are all these systems really turbulent after an indefinitely long time? Do they behave completely differently, or do they share

common behaviours, beyond the superficial fact that they all have topologically similar defects? Can defect-mediated turbulence perhaps be controlled? Much research is now being focused on these questions, and pieces of the puzzle are starting to be assembled. □

Alain Karma is in the Department of Physics and Center for Interdisciplinary Research on Complex Systems, Northeastern University, Boston, Massachusetts 02115, USA.

OBITUARY

Stanley Keith Runcorn (1922–95)

THE death of Keith Runcorn, who was murdered in San Diego on 5 December, has shocked geophysicists the world over. Runcorn was a central player in two of the crucial debates in Earth science in the mid-twentieth century — those over the origin of the Earth's magnetic field and the validity of the theory of continental drift. At the time of his death he was on his way to a meeting from the University of Alaska, where he had occupied an endowed chair following his retirement in 1988 from the University of Newcastle upon Tyne.

Lancashire born, Runcorn read mechanical sciences at Cambridge during the Second World War, and then planned to do research on cosmic rays under Patrick Blackett at Manchester. He became a geophysicist almost by accident, by becoming involved in a long-standing controversy over the Earth's geomagnetic field. Blackett was developing a theory to explain the origin of the field, and in 1947 published it in a *Nature* paper entitled "The magnetic field of massive rotating bodies". These ideas were jelled by the addition of H. Babcock's observation that the ratio between the magnetic moment and angular momentum for the star 78 Virginis is similar to that of the Earth, and also to that previously determined for the Sun. Thus magnetic moment was inferred as a fundamental property of all massive rotating bodies. If this theory were correct, the geomagnetic source region must comprise the whole Earth out to its surface. On the other hand, the geodynamo theory of Walter Elsasser and Edward Bullard required the field to originate entirely within the Earth's outer core of liquid iron.

Runcorn decided that it should be possible to distinguish between the two ideas by taking magnetometers down the coal mines of his native Lancashire to measure the variation of geomagnetic field intensity with depth. The result turned out to favour the 'core' principle. Two of the students who worked on the

experiment, Raymond Hide and Frank Lowes, subsequently gained eminence for their own contributions to geophysics. Others were to follow.

Blackett had preferred to test his rotating body theory by detecting whether a weak magnetic field existed near a rotating gold sphere, and he designed a highly sensitive astatic magnetometer for this purpose. The result was negative, and he sought another



Godfrey Argent

use for his magnetometer. With John Clegg, he found that it could measure weak remanent magnetization in sedimentary rocks, and thereafter he became interested in rock magnetism.

In 1950, Runcorn moved to Cambridge and started a research group in palaeomagnetism; Ted Irving and myself were his first research students. The team grew and within a few years showed that the British Isles had been moving relative to the geographical poles throughout the Phanerozoic; the resulting 'polar wander' curve attracted wide public as well as scientific attention. In 1955 Runcorn moved to Newcastle, and after several summers' field work produced a 'polar wander' curve for North America.

Irving, working in Canberra, carried out a similar survey for Australia, and from Newcastle I did the same for South America. Alan Nairn worked on rocks from Africa. Meanwhile the Blackett/Clegg group, now in London, worked mainly on India. The upshot was that it became clear that the continents had moved horizontally relative to each other, as well as to the geographical poles, broadly verifying Alfred Wegener's concept of continental drift. But the palaeomagnetic evidence was regarded with suspicion by many geologists, and it took almost two decades for this revolution in the Earth sciences to become generally accepted, even when the evidence from sea-floor spreading and the theory of plate tectonics came along.

Keith Runcorn remained unmarried, which gave him the freedom to devote his life to his science. He was a restless traveller, so much so that at Newcastle he became fondly known as the 'visiting' professor of physics. He took such remarks in good part because he knew, as we all did, that perhaps more than any other person it was he who put the university on the map scientifically. On the other hand, he never showed any interest in university politics, nor did he become much involved in the administration of scientific societies. He did, however, have a vision of a pan-European society for geophysics, equivalent to the American Geophysical Union. This vision has become a reality in the form of the flourishing European Geophysical Society.

For all his achievements and many awards, Keith Runcorn remained modest and unpretentious. His reputation as one of the foremost geophysicists of his generation will remain long after his death, as will memories of his warmth and sincerity.

Kenneth M. Creer

Kenneth M. Creer is in the Department of Geology and Geophysics, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JW, UK.

Tyrannosaurus on the run

R. McNeill Alexander

ONE of the characters in *Jurassic Park*¹ tells us that *Tyrannosaurus rex* could easily outrun a Jeep, driven off-road at 30–40 m.p.h. (13–18 m s⁻¹), and the film of the book showed us how alarming that would have been. The suggested speed is similar to that of racehorses (16–17 m s⁻¹, based on race times given in newspapers). In a paper just published in *Journal of Vertebrate Paleontology*, however, Farlow *et al.*² argue not only that the top speed of *Tyrannosaurus* was much slower (about 10 m s⁻¹), but that there were reasons why the creature would have found it dangerous to run at the higher speeds.

Old reconstructions of the giant carnivore show an ungainly and presumably slow-moving animal. Bakker's³ arguments for dinosaurs being warm-blooded, and the stunning pictures he drew to illustrate them, encouraged us to think of dinosaurs as much faster and more athletic, but an analysis of bone strength⁴ seemed to show that *Tyrannosaurus* must have been relatively slow. The faster an animal runs or the higher it jumps, the greater the forces on its feet. If similar animals (possibly of very different sizes) move in dynamically similar fashion, the forces on their feet will be equal multiples of body weight. The dimensions of fossil leg bones can be used to estimate the strengths of the living bones, and if these are equal multiples of body weight, for a fossil species and a living one, the fossil species could have been as athletic as the living one. This approach led to the conclusion that *Tyrannosaurus* probably ran only slowly, with a top speed possibly around 8 m s⁻¹. I even claimed (I was younger then) to be able to outrun *Tyrannosaurus*.

Farlow *et al.*² have repeated this analysis with better material — a skeleton of *Tyrannosaurus* in the Museum of the Rockies — and have reached the same conclusion. But they also present a new argument for *Tyrannosaurus* being slow. The animal was tall, with the lowest points on its belly and head being 1.5 and 3.5 m, respectively, above the ground. If it tripped it would fall a long way, and its vestigial arms would be useless to break its fall. It could not risk falling, runs the argument, and would have had to have moved carefully.

From simple calculations, Farlow *et al.* estimated impact forces occurring during falls. If the torso falls from 1.5 m and is brought to rest in 0.3 m (an estimate that includes deformation both of the wall of the chest, and of the ground), the mean force during the impact is $(1.5 + 0.3)/0.3 = 6$ times torso weight. If the animal falls while running at 20 m s⁻¹ and skids 3 metres along the ground before coming to a halt, the horizontal force involved is its kinetic energy divided by the skid distance, 7 times

body weight. The deformation distance and the skid distance are merely plausible guesses; even so, Farlow *et al.* conclude that the dinosaur would have been unlikely to survive its fall.

There is very little quantitative information about the forces involved in falls and similar accidents either for animals or for people. Indeed, I know of only one sporting accident for which forces could be calculated: a weight lifter broke his patellar ligament in competition, while being filmed by a team of biomechanists who were able to calculate that the force that broke it had been 14,500 newtons (ref. 5). We might be able to learn more by fitting appropriate instruments to footballers and sending them out to play, hoping they would get hurt, but that would seem callous. However, we do know a good deal about the forces that cause injury in car crashes⁶. It proved possible to determine the forces that had acted on seat belts of one particular design, by examining the belts after accidents, and it was found that young adults suffered no chest injuries unless belt force exceeded 7,300 newtons (about 10 times body weight for a typical man). To estimate the force on the chest we must take account of the belt being attached at both ends, and of the angles of its attachments: 15 times body weight seems likely.

The forces that animals of different sizes can withstand are not expected to be proportional to their body weight, but to their areas (to the two-thirds power of weight). If a 70-kg man can tolerate 15

times body weight, a 6,000-kg tyrannosaurus can be expected to tolerate $15 \times (70/6,000)^{1/3} = 3.4$ times body weight. So according to these calculations it does indeed seem that if a tyrannosaurus fell while going full pelt it would cause itself some damage.

However, we need to ask whether animals can be expected to be cautious, or to live dangerously. The underside of a giraffe's belly is about 2.0 m from the ground (compared to 1.5 m for *Tyrannosaurus*), but giraffes gallop moderately fast, up to at least 11 m s⁻¹ (ref. 7). Ostriches are bipeds with no arms to break a fall, but run very fast indeed (I have driven alongside one in a Jeep with the speedometer reading 35 m.p.h., 16 m s⁻¹). And the argument that *Tyrannosaurus* must have moved cautiously because of the danger of falling would seem to imply that, for instance, gibbons should move gingerly through the treetops instead of swinging rapidly from their arms.

Whatever the merits of the new argument of Farlow *et al.*, the old one based on leg-bone strength is confirmed. *Tyrannosaurus* was not good at chasing Jeeps. □

R. McNeill Alexander is in the Department of Pure and Applied Biology, University of Leeds, Leeds LS2 9JT, UK.

1. Crichton, M. *Jurassic Park* (Century, London, 1991).
2. Farlow, J. O., Smith, M. B. & Robinson, J. M. *J. vert. Paleontol.* **15**, 713–725 (1995).
3. Bakker, R. T. *Nature* **238**, 81–85 (1972).
4. Alexander, R. McN. *Zool. J. Linn. Soc.* **83**, 1–25 (1985).
5. Zernicke, R. F., Garhammer, J. & Jobe, F. W. *J. Bone Jt Surg.* **59A**, 179–183 (1977).
6. Society of Automotive Engineers SAE Handbook Section 34, 224–241 (Soc. Auto. Engs. Warrendale, Pennsylvania, 1989).
7. Alexander, R. McN., Langman, V. A. & Jayes, A. S. *J. Zool.* **183**, 291–300 (1977).

PARTICLE PHYSICS

Glueball caught in a lattice

Frank Close

ATOMS of light do not exist, but theory suggests that closely related entities called 'glueballs' should occur on the scale of individual protons and neutrons. For over twenty years no clear evidence for glueballs has emerged, though there have been occasional sightings whose status is still not settled. In the past year this question has come sharply into focus, culminating in the most recent "Numerical evidence for a scalar glueball"¹. One press report described this as the "first instance of a particle's discovery by computer".

First, what are glueballs and how do they relate to atoms of light?

Many of our daily experiences are governed by the absorption and emission of photons, the quanta of the electromagnetic field and thus the particles that carry electromagnetic forces. An essential

rule underpinning the structure of matter is that opposite charges attract, and it is the attraction between electrons and protons that allows atoms to form. Photons carry no electric charge and so do not attract one another, hence the absence of atoms of light.

The relativistic quantum theory of electromagnetic phenomena is known as quantum electrodynamics (QED) and is considered a model for theories of the other fundamental forces, such as those controlling the structure of the atomic nucleus and its constituent particles. Quantum chromodynamics (QCD) is the theory of the strong nuclear force, and as its name suggests, it is tantalizingly similar to QED. The analogues of photons are 'gluons'. Whereas photons carry no electric charge, gluons do possess the 'colour'

charge that is the source of the strong force. So although photons do not clump into 'light-balls', gluons should form short-lived clusters, whimsically deemed 'glueballs'.

The similarity between QED and QCD encourages attempts to construct grand unified theories of matter and forces. Whereas the existence of gluons has been established in experiments involving collisions of particles at extremely high energies, glueballs remain a missing link in the Standard Model of particle physics, largely because their properties are so hard to predict.

Analytic solutions to the equations of QCD, which could definitively establish the properties of glueballs, are still beyond us. The best simulations have been those in which space and time are treated as a discrete lattice of points, and calculations are made numerically by computer; this is known as 'lattice QCD'. By 1993, calculations by the UKQCD collaboration² predicted that the lightest glueball would be a 'scalar' (have no intrinsic spin) with a mass between 1500 and 1600 MeV c^{-2} (about 60% more than the proton). A scalar particle of mass 1500 MeV c^{-2} was discovered at CERN last year, and studies of its properties have added to the suspicion that this either is, or is closely related to, the scalar glueball^{3,4}. Now we have the largest computer study to date¹, which reports a mass of around 1700 MeV c^{-2} . This is near enough to 1500 that, bearing in mind the approximations in the computation, it could be further support for the CERN particle's being the glueball. However, the new computation, from IBM, goes further than before and, in this, provides a tantalizing twist.

Glueballs are expected to be unstable, decaying into well-known particles such as pions, kaons or etas in less than 10^{-23} seconds. Hitherto, predictions for the absolute lifetime of glueballs and the relative abundance of these various particles in their decay debris have been poorly constrained. This is where the results from IBM are novel and will stimulate debate — for the first time a computer solution to the underlying theory (albeit with some technical simplifications to make the calculation practicable) predicts a mass, a lifetime and a pattern of decay that are tantalizingly close to those of a well-established particle known as the 'theta', of mass 1710 MeV c^{-2} . It is this that has encouraged the authors to claim "strong evidence that [the theta] is largely a scalar glueball" and that the 1500 MeV c^{-2} particle is a conventional particle consisting of a quark and antiquark bound together.

During the next year further detailed studies of the '1500' particle will continue at LEAR (the CERN Low Energy Antiproton Ring). As concerns the theta, the '1710' candidate, there are vital ques-

tions that need to be settled. First, it is necessary to establish whether it is even a scalar particle, as required of the glueball predicted by IBM. Many experimentalists suspect that the theta is actually an unresolved mix of two particles, one of which is a scalar and the other a tensor (having an intrinsic spin of two). If this is the case then the tensor and scalar contributions to the composition of the decay products will also need to be resolved before the jury delivers its verdict. The current 'official' table of particle properties⁵ is not prepared to deliver judgement on the spin of the theta, saying only that it is either scalar or tensor; and as for the decay products it lists the various particles only as 'seen', with no quantitative measurements.

It is ironic that there may be two scalar particles with masses so near to the predictions emerging from two separate lattice QCD computations. The glueball question is now shifting from one of existence to a judgement of Solomon on

which of two claims is the reality. If the theta is eventually established as a scalar with the properties predicted by the lattice computation, that will be highly significant. Not only will it confirm that the first glueball has been found, and so signpost how best to explore this new spectrum of states, but it will prove the usefulness of lattice computer calculations. That would have profound implications for particle physics, where lattice QCD is already a major tool in analysis, and perhaps for the wider field of mathematical physics. □

Frank Close is in the Theoretical Particle Physics Division at the Rutherford-Appleton Laboratory, Oxfordshire OX11 0QX, UK.

1. Sexton, J. et al. *Phys. Rev. Lett.* **75**, 4563–4566 (1995).
2. Bali, G. et al. *Phys. Lett.* **B309**, 378–382 (1993).
3. Amsler, C. & Close, F. E. *Phys. Lett.* **B353**, 385–390 (1995).
4. Amsler, C. & Close, F. E. *Phys. Rev. D* (in the press).
5. Particle Data Group *Phys. Rev. D* **50**, 1173–1826 (1994).

PALAEOANTHROPOLOGY

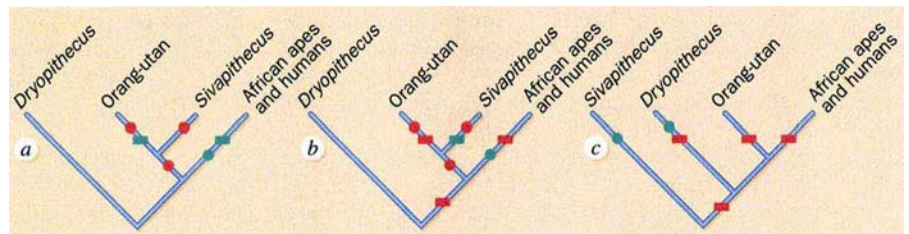
The nature of the evidence

Peter Andrews and David Pilbeam

THE systematics of fossil mammals is largely dependent on analysis of teeth and skull bones. Even when bones from the rest of the skeleton are available, all too often they are regarded of interest only in functional terms and are not accorded any value in analyses of evolu-

very clearly. The fossil is discussed by Moyà-Solà and Köhler on page 156 of this issue².

Dryopithecus was first described three years before Darwin's *Origin of Species* was published, and was the first known fossil ape. The type species, *D. fontani*, is



Possible relationships of the fossil apes *Dryopithecus* and *Sivapithecus* with the living great apes, depending on the interpretation of cranial and postcranial characters. *a*, *Sivapithecus* could be related to the orang-utan on the basis of cranial characters (red circles), and postcranial characters would then be identified as independently derived, convergent characters (homoplasies; green squares) in the living apes. *b*, As *a*, but if the postcranial characters shared by the living apes are homologous, that is, of common evolutionary origin (red squares), *Sivapithecus* must have a unique reversal in these characters (green square). *c*, Postcranial characters are again interpreted as homologous in the living apes, together with *Dryopithecus* (red squares), in which case the same characters in *Sivapithecus* would be primitive and the cranial characters linking it with the orang-utan would be primitive or homoplasies (green circles).

tionary history¹. More postcrania are now being found as a result of improved excavation and collecting techniques, and the patterns of evolutionary change that they show have to be integrated with those evident from the jaws and teeth. This is not always a straightforward matter, as the discovery of a partial skeleton of *Dryopithecus*, a fossil ape from Spain, shows

based on a lower jaw with most of its teeth in place; the shaft of a humerus was found later³ at the same site in the south of France. This gave an early indication that at least this part of the postcranial skeleton was more similar to the corresponding parts of living great apes than were the jaws and teeth, which retained many characters unchanged from the

ancestral hominoid form⁴. This divergence in evidence between teeth and postcrania became greater when additional dryopithecine postcrania from Rudabanya in Hungary were analysed, for these also were shown to have greater similarities with living great apes than other fossil apes⁵.

The postcranial skeleton found by Moyà-Solà and Köhler² is some 9.5 million years old, and comes from an extension of a well-known dryopithecine site in the Valles Penedes region of Spain. A partial skull was described from this site by the same authors^{6,7}, who agreed with the present consensus view that it should be assigned to the species *D. laietanus*. There is some controversy, however, over the phyletic affinities of this material⁷⁻⁹, with current proposals including a relationship with the African ape and human clade¹⁰, with the orang-utan^{6,7}, or with neither^{4,11}, and this controversy is likely to be increased rather than diminished by the recent discoveries.

The interpretation of the postcranial characters as shared, derived characters of the *Dryopithecus* plus African ape and human clade¹⁰ has little support from cranial characters. Support for its relationship with the orang-utan is also equivocal, whereas there is strong cranial evidence that the Eurasian genus *Sivapithecus* belongs to the orang-utan clade^{12,13} while retaining most aspects of the ancestral postcranial conditions¹⁴. These two sources of evidence are incompatible for the two fossil hominoids, but there are several ways of interpreting them (see figure).

First, one would claim that if the evidence of the shared characters of the skull in *Sivapithecus* and the orang-utan is accepted, the postcranial characters shared by the orang-utan and the other apes should be interpreted as independently derived, convergent characters (homoplasies), for they are not present in *Sivapithecus*. Second, it is possible that these characters could have been present in the common ancestor of the great apes if the condition in *Sivapithecus* represents a unique reversal. Third, if the postcranial characters shared by the great apes and *Dryopithecus* are homologous, *Sivapithecus* could be excluded from any part of that clade and the cranial characters linking it to the orang-utan would have to be homoplasies. Finally, it is again possible that the postcranial characters are homologous and that the cranial features shared by *Sivapithecus* and the orang-utan are primitive.

There is no reason to expect all parts of the body to change at the same rate and at the same time, and this is clearly not the case in various hominoids from the Miocene (some 8–15 million years ago). The functional interpretation of the bones of the skeleton at least is clear:

they show some adaptations to suspension that are similar to those seen today in the orang-utan. The implications of this are that *Dryopithecus* was a below-branch arboreal quadruped with a form of locomotion similar in many respects to that of the orang-utan (although the fossil ape was smaller). Most other fossil apes retained a form of locomotion^{4,5} which is similar to that of present-day Old World monkeys and which is interpreted as ancestral for the hominoid primates. This change is manifested in the presence of several characters shared by *Dryopithecus* and the extant great apes and humans.

The question remains, however, as to which set of characters to use as the basis for phylogenetic interpretation, those from the face or those from the postcranium. Combining them in a total evidence cladistic analysis¹⁵ does not solve the problem, for such an analysis would be heavily dependent on which morphological area can muster the greater number of characters. Functional significance might provide some indication of character independence, but it may be questioned if the functional complexity of the skull, though poorly known, renders it better or worse than the relative functional simplicity of the postcranial elements, be they ever so well known. Character complexity, phenotypic or genetic or both, is another criterion.

Additional fossil specimens always help, although they too can add to the problem, for such discoveries all too often lead to more questions than are answered and even greater uncertainty about evolutionary pathways. Arising out of the considerable incongruence between supposed shared-derived similarities revealed here, we believe that parallel work must concentrate on the way characters are initially chosen for phylogenetic analysis. □

Peter Andrews is in the Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. David Pilbeam is in the Department of Anthropology, Harvard Peabody Museum, 11 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

1. Novacek, M. J. *Nature* **356**, 121–125 (1992).
2. Moyà-Solà, S. & Köhler, M. *Nature* **379**, 156–159 (1996).
3. Pilbeam, D. & Simons, E. L. *Nature* **229**, 408–409 (1971).
4. Benefit, B. R. & McCrossin, M. L. A. *Rev. Anthropol.* **24**, 237–256 (1995).
5. Begun, D. *Am. J. phys. Anthropol.* **87**, 311–340 (1992).
6. Moyà-Solà, S. & Köhler, M. *Nature* **365**, 543–545 (1993).
7. Moyà-Solà, S. & Köhler, M. *J. hum. Evol.* **29**, 101–139 (1995).
8. Begun, E. R. *J. hum. Evol.* **29**, 169–180 (1995).
9. Martin, L. & Andrews, P. *Nature* **365**, 494 (1993).
10. Begun, D. *Yb. Phys. Anthropol.* **37**, 11–63 (1994).
11. Andrews, P. *Nature* **360**, 641–646 (1992).
12. Pilbeam, D. *Nature* **295**, 232–234 (1982).
13. Andrews, P. & Cronin, J. *Nature* **297**, 541–546 (1982).
14. Pilbeam, D., Rose, M., Barry, J. & Shah, I. *Nature* **348**, 237–239 (1990).
15. Smith, A. B. *Systematics and the Fossil Record* (Blackwell Science, 1994).

DAEDALUS

The light metallic

EVEN a perfectly transparent substance reflects some light from its surface. If, like diamond, it has a high refractive index, it reflects quite a lot. Now metals have extremely high refractive indices, so Daedalus suspects that they too are perfectly transparent. They just reflect away all the light that hits them, so that none can get in. By the same token, of course, none can get out either. Light inside a metal would be perfectly internally reflected, and could never emerge.

How to generate light inside a metal? Laser action is the obvious way. Some metals can absorb molecules that lase quite readily. Thus iron absorbs carbon monoxide, and palladium absorbs hydrogen and deuterium. Musing on this, Daedalus recalled 'cold fusion', claimed to occur when deuterium was electrolytically injected into palladium electrodes. He reckons the deuterium arrived in a highly excited state, and lased its energy away along unexpected metallic channels. The resulting heat was misinterpreted as fusion energy.

So DREADCO's physicists are devising ways of injecting suitably excited molecules into solid metals. They are trying electrolysis under vast voltages, ion bombardment and energetic surface decompositions. The metal laser is welded to a long wire, a sort of metallic optic fibre down which the laser beam will escape. Its far end, coated with caesium metal, is in a vacuum chamber. A photon hitting a caesium surface, of course, can eject an electron into a surrounding vacuum: this is the classic photoelectric effect. Daedalus reckons it will work just as well if the photon hits the caesium surface from inside.

Once intra-metallic radiation can be generated and detected, metal laser and light-wire technology will transform communications. The optic fibre companies will lose their monopoly of wide-band optical transmission: their telephone rivals will suddenly be able to relay optical signals over their networks of antique copper wire. They in turn will be upstaged by the electricity companies, who will fire wide-band optical data down their power lines, while the more historic gas and water companies join the game on their lead pipes. Endless television, video and computer data will shuttle in and out of every home along the mass of new fast data lanes. Sadly, they won't be all that fast. A light-wire, with its high refractive index, will inevitably act as a delay line. Quickfire repartee will arrive annoyingly late, and all parties to a video-conference will find themselves waiting impatiently for the others to see the point.

David Jones

Use of walking trails by bees

SIR — Trail use has evolved twice among the ground-dwelling social insects (ants and termites), a striking example of evolutionary convergence in behaviour. The other groups of social insects, bees and wasps, are principally flying organisms, not known to use extended walking trails for locating colony resources, although some species of stingless bees follow scent marks or aerial odour trails during flight¹. Our field investigation of the Amazonian bumble bee, *Bombus transversalis*, reveals that it clears and maintains trails on the forest floor, similar in appearance to the recruitment trails of ants².

Most bumble bees inhabit the temperate areas of the world, occurring as far north as



Bombus transversalis workers following one another in tandem along a trail. The field study was carried out in the Tambopata Nature Reserve³, located along the Rio Tambopata (12° 50' S; 69° 17' W), Department of Madre de Dios, Peru, approximately 280 m above sea level, during the dry season May–June 1995. Colony located by T. Ruiston and P. Debes.

the Arctic Circle and at elevations up to 5,000 m. Only *B. transversalis* is known to inhabit the tropical forests of the Amazon Basin³. Our field observations focused on a colony in its relatively early growth phase, containing around 350 workers and a single queen. Specialized workers foraged for nectar and pollen throughout the daylight hours. Another subset (behavioural caste) of workers devoted much time to building and maintaining an elaborate nest canopy⁴ composed of a stiff mass of leaves, rootlets and fibres woven into an aerated cone 10–15 cm thick. The use of marked bees indicated that this subset of workers (8–10% of colony adults) was involved in canopy building for several days at least.

Examination of how the materials for the nest canopy were gathered revealed that 'thatch-makers' walked off the nest cone periodically and followed one another in tandem along one of two cleared paths on the forest floor (see figure). Extending from opposite sides of the nest, the paths were 9–10 cm wide and 1.5 m and 2 m long, respectively. Both paths were cleared of vegetation, and terminated beneath a dense mat of fallen leaves, where bees

often remained hidden for more than 5 minutes. Bees walking outbound along the paths interrupted their movement periodically to jab at the soil surface with their mandibles, and to cut and remove or investigate plant fragments they encountered. Bees emerging from the terminal leaf litter flew directly back to the nest cone, or walked back along the cleared path and proceeded to work on the canopy.

The trails seem to be actively maintained, for when we dropped fragments of leaf litter and twigs along selected portions of the front and rear paths, outbound thatch-makers pushed them sideways off the path or backwards towards the nest cone. Items were removed within 5–10 s by the first bees encountering them, and most items were detected within 5 minutes. Observations of tagged individuals indicated that 23 out of 25 bees remained on their respective trails (92% trail specificity).

Six weeks after our initial observations, the original two trails had been extended by a further 0.5 m and three more had been newly cleared (1.5–2 m long; M. Cohen and N. Thorp, unpublished observations). The bees appeared to minimize the overlap in collection of leaf litter on the forest floor by building the

first two trails on opposite sides of the nest (180°), adding subsequent trails at maximally distant positions relative to the first (90° and 45°, respectively). The canopy increased in height and width by 10 cm during the same 6-week period. Estimated collection rate of thatching material is about 12 cm² per day.

The paradoxical use of terrestrial trails by a bee that collects its food resources by flying may, in fact, be an efficient response to the challenges this species faces in nest insulation and defence. The use of trails enables the colony to encompass a wider and more rapidly defended home range, and facilitates the efficient collection of nest-building materials.

S. A. Cameron*, J. B. Whitfield†

Departments of *Biological Sciences
and †Entomology,
University of Arkansas,
Fayetteville, Arkansas 72701, USA

1. Wille, A. A. *Rev. Ent.* **28**, 41–64 (1983).
2. Hölldobler, B. & Wilson, E. O. *The Ants* (Harvard Univ. Press, Cambridge, MA, 1990).
3. Milliron, H. E. *Mem. ent. Soc. Can.* **89**, 81–237 (1973).
4. Olesen, J. M. *J. trop. Ecol.* **5**, 243–246 (1989).
5. Pearson, D. L. & Dressler, R. L. *J. trop. Ecol.* **1**, 37–54 (1985).

Robustness of cooperation

SIR — The 'Prisoner's Dilemma' is used extensively to study the conflict between individual and collective rationality^{1,2}. Two players can either cooperate, *C*, or defect, *D*. There are four possible pay-offs: *R* each, when both players cooperate; *P* each, when both players defect; or *T* to the defector and *S* to the cooperator in the remaining cases (usually $T > R > P > S$ and $2R > T + S$). Nowak and May extended the game to a population of players arranged in a closed two-dimensional $n \times n$ lattice. In each period, each player plays the Prisoner's Dilemma with its eight immediate neighbours and itself, adopting the strategy used by its opponent with the highest score from the previous period. Computer simulations using the pay-offs $2 > T > 1.8$, $R = 1$ and $S = P = 0$ show that the asymptotic fraction of *C* players, f_C , fluctuates around 0.318. Recent articles have used this spatial version of the Prisoner's Dilemma to explain how "cooperation rather than exploitation can dominate in the darwinian struggle for survival"^{3–5}.

We have analysed three independent natural variations of the spatial Prisoner's Dilemma. We find that cooperation is eliminated in all three, suggesting that the spatial Prisoner's Dilemma cannot fully account for the emergence and persistence of cooperation in natural and social systems. First (Fig. 1a), we considered the effect of players making independent and identically distributed errors. With a small probability ϵ , each player errs and chooses evenly between strategies *C* and *D*; with probability $1 - \epsilon$, the player follows the Nowak and May update rule. Second (Fig. 1b), we considered the effects of different levels of synchronization as opposed to the condition of complete asynchrony⁶. During each period, players fail to update their previous strategy with a small probability, θ . Third (Fig. 1c), we considered the effect of a small percentage of the cooperators resorting to cheating in order to exploit their neighbours. After following the Nowak and May update rule, each cooperator has a small independent probability, ϕ , of cheating by switching to defection.

In all three variations, cooperation did not persist. Previous reports^{7,8} on probabilistic perturbations are based on probabilistic ratios of the pay-offs, and can only affect the boundary between clusters of cooperators and defectors. Our variations allow all players to commit errors with a low probability and are not restricted to such boundaries. A lone defector can grow into a cluster of

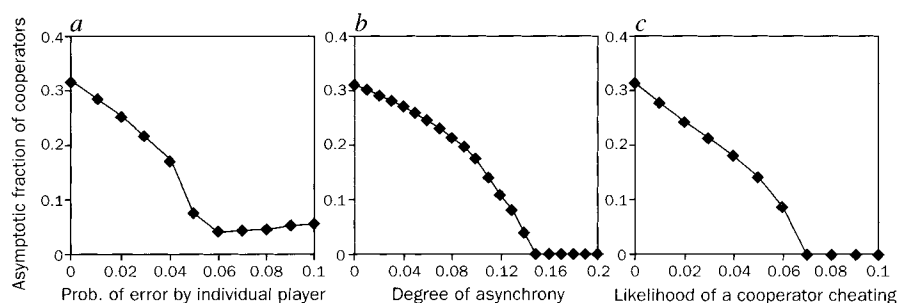


FIG. 1 Simulations were performed on a 100×100 lattice, for 500 periods starting with a random initial configuration with 10% defectors (and 90% cooperators) and pay-offs $T = 1.85$, $R = 1$, $P = S = 0$. *a*, The asymptotic fraction of cooperators, f_c , decreases rapidly with small increases in the error factor ϵ (the probability with which each player errs and chooses evenly between C and D). As ϵ approaches 0.06, some isolated cooperators remain, but they are too few and too dispersed to form clusters of cooperation. With further increases in ϵ , the number of cooperators increases, because $\epsilon/2$ of the population will randomly choose to cooperate. However, no clusters of cooperation were observed. *b*, A high degree of synchronization is required to maintain cooperation. Each player was allowed to skip updating its strategy during a period with a small independent probability, θ . With $\theta = 0$ (100% synchronization), $f_c \approx 0.318$. When θ was increased to 0.1, f_c dropped by half. Cooperation is eliminated when θ reaches 0.15. Thus, a relatively small percentage (~15%) of the population not being synchronized, as opposed to the condition of complete asynchrony⁶, is sufficient to eliminate cooperation. *c*, Cooperation is rapidly reduced if each cooperator has a small independent probability, ϕ , of cheating after following the Nowak and May update rule. The fraction of cooperators, f_c , quickly decreases as ϕ increases. Clusters of cooperation are wiped out when about 7% of the cooperators cheat.

defectors. Clusters of cooperators and defectors grow and diminish, producing a variety of spatial patterns. The asymptotic behaviour shows small, tightly knit clusters of cooperators gliding around in a world of defectors. However, even one member of a cluster of cooperators changing to defection can eliminate or shrink the cluster. For cooperation to persist in the spatial Prisoner's Dilemma, a high degree of synchronization with no errors is required, a condition that is unlikely in most natural and social systems.

Arijit Mukherji

Carlson School of Management,
University of Minnesota,

Vijay Rajan

James R. Slagle

Department of Computer Science,
University of Minnesota,
Minneapolis, Minnesota 55455, USA

NOWAK *ET AL.* REPLY — Mukherji *et al.* investigate the question of whether cooperation can survive in the spatial Prisoner's Dilemma in the presence of noise. They perform computer simulations of three different types of stochastic perturbation: (1) a fraction of sites is occupied at random by cooperators or defectors; (2) a fraction of sites is not updated, but remains with the current strategy; and (3) a fraction of cooperators turns spontaneously into defectors (this third assumption is well chosen for attempting to eliminate cooperators). All their simulations are restricted to the particular parameter region $2 > T > 1.8$.

We have explored Mukherji *et al.*'s extensions of our model, but for the wider range of parameters outlined in our original papers. For the particular region considered by Mukherji *et al.*, we of course

confirm their results, but the spatial Prisoner's Dilemma has nine different parameter regions for $2 > T > 1$ (see refs 3, 4, 7–9). For the other parameter regions, we find in all three cases that cooperators can persist despite significant amounts of noise (see Fig. 2). In model (1) the baseline level of cooperators is $\epsilon/2$ (because of the random assignment of a fraction ϵ of sites). We find that in parameter regions 1–5, cooperators exist at abundances well above baseline for noise levels up to 25%, and in parameter regions 1–3, for noise levels up to 50% and more. In model (2) the degree of asynchrony has essentially no effect on the asymptotic abundance of cooperators in parameter regions 1–8. In model (3), survival of cooperators is possible in parameter regions 1–6 for noise levels of about 13% and in regions 1–3 for noise levels of about 26%.

The main conclusion in our original papers^{3,7–9} was that spatial structures can facilitate the survival of cooperators. It is clear that any kind of noise that tends to destroy spatial structures will work against the survival of cooperators. But, even in the region $2 > T > 1.8$, our original results for the deterministic case remain valid at low noise levels and are lost only when the noise exceeds a threshold magnitude. Our present simulations (and previous work^{7,8,10}) also show that the spatial Prisoner's Dilemma — if seen in the whole range of parameter regions — is, in fact, robust against significant amounts of stochastic perturbations, as well as several other complications⁸.

In contrast with Mukherji *et al.*'s final conclusion, we contend that cooperation persists in the spatial Prisoner's Dilemma even in the face of reasonably high levels of noise.

Martin A. Nowak

Sebastian Bonhoeffer

Robert M. May

Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK

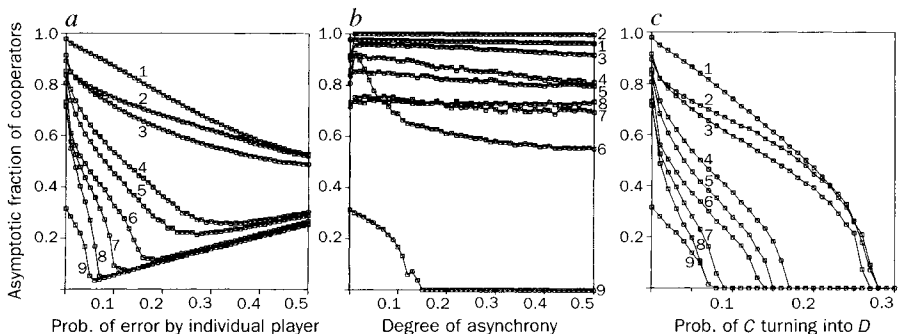


FIG. 2 Computer simulations for the spatial Prisoner's Dilemma with three different types of stochastic perturbations as proposed by Mukherji *et al.*, but for all 9 relevant parameter regions of the game^{7,8} represented by the T values 1.05, 1.13, 1.16, 1.35, 1.42, 1.55, 1.71, 1.77 and 1.9 (corresponding to labels 1–9 in the figure). Mukherji *et al.* show results for the ninth parameter region $2 > T > 1.8$. The x-axis denotes: *a*, the fraction of sites given randomly to cooperators or defectors in each generation; *b*, the fraction of sites that are not updated in every generation; *c*, the fraction of cooperators that are changed into defectors in every generation. The y-axis is the asymptotic fraction of cooperators in a 100×100 array after 500 generations with an initial condition of 90% cooperators.

1. Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
2. Rapoport, A. in *The New Palgrave: A Dictionary of Economics* Vol. 3, 973–976 (Macmillan, London, 1987).
3. Nowak, M. A. & May, R. M. *Nature* **359**, 826–829 (1992).
4. Nowak, M. A., May, R. M. & Sigmund, K. *Scient. Am.* **272** (6), 76–81 (1995).
5. Sigmund, K. *Nature* **359**, 774 (1992).
6. Huberman, B. A. & Glance, N. S. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7716–7718 (1993).
7. Nowak, M. A., Bonhoeffer, S. & May, R. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4877–4881 (1994).
8. Nowak, M. A., Bonhoeffer, S. & May, R. M. *Int. J. Bifurcation Chaos* **4**, 33–56 (1994).
9. Nowak, M. A. & May, R. M. *Int. J. Bifurcation Chaos* **3**, 3578 (1993).
10. Herz, A. V. M. *J. theor. Biol.* **169**, 65–87 (1994).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words.

Through the looking glass

Martin Gardner

Lewis Carroll: A Biography. By Morton N. Cohen. Knopf/Macmillan: 1995. Pp. 577. \$35, £25.

THE most surprising of many surprises in Morton Cohen's splendid new biography of the Reverend Charles Lutwidge Dodgson, Lewis Carroll's real name, is the conjecture that Dodgson not only was romantically in love with Alice Liddell, 20 years his junior, but also for a time actually hoped he might marry her. Not that he ever formally proposed. More probably, Cohen believes, he approached Alice's parents with a suggestion that "perhaps in the future, if her [Alice's] affection for him did not diminish, he would be happy to propose".

We do know that Mrs Liddell, wife of the Dean of Christ Church, Oxford, where Carroll taught mathematics, was suddenly impelled to burn all of Dodgson's letters to Alice and to forbid him for a time to see her three charming daughters. The pages in Dodgson's carefully kept diary covering this painful period were destroyed after his death by his niece. And there were rumours in Oxford, as one letter-writer put it, "that Dodgson has half gone out of his mind in consequence of having been refused by the real Alice".

There has been much speculation about Dodgson's passion for photographing naked little girls. Although he eventually tried to destroy all these photos — there had been unpleasant gossip about them — four pictures survived. Cohen reproduces all four. One, a full-frontal view of a beautiful little girl resting supine, would surely be rejected today by *Playboy* as too close to child pornography. Dodgson insisted that such photography was a pure, untarnished effort to capture the beauty and innocence of pre-pubescent female bodies. Is it possible he deceived himself?

Cohen thinks he did. In a remarkable chapter titled "The Fire Within", he argues that Dodgson had all the desires of a heterosexual man, but because of his Anglican faith, augmented by Victorian mores, he struggled all his life to conceal and repress such feelings. Cohen believes it was the frequency of erotic images in Dodgson's dreams and reveries that explains the many passages in his diary where he seeks God's help in overcoming what he considered to be deplorable sins.

No one knows more about Dodgson than Professor Cohen, now retired from teaching English literature at City College of New York. He has edited two volumes of Dodgson's letters and a book of essays about Dodgson. His skill and tact in covering all aspects of Dodgson's complex, enigmatic personality is impressive. The



Charles Lutwidge Dodgson: his love for mathematics and logic was unbounded.

book is no hagiography. It covers all Dodgson's foibles: his prudishness, his snobbishness, his fastidiousness, his amusing eccentricities and his indifference to literature and events outside England. All this is overshadowed, of course, by his immense creative powers.

Many aspects of Dodgson's writings and beliefs are of special interest to readers of *Nature*. Dodgson was not a great mathematician, but his love for mathematics and logic was unbounded. One reason the *Alice* books are so relished by mathematicians is because they bristle with mathematical riddles, paradoxes, jokes and witty wordplay. Two of his books are about mathematical puzzles: *A Tangled Tale* and *Pillow Problems*. More serious mathematical works include *Euclid and His Modern Rivals* and *An Elementary Treatise on Determinants*. His interest in

formal logic produced *Symbolic Logic* and *The Game of Logic*. In recent years the discovery of Dodgson's unpublished papers on logic show him to have been well ahead of his time in recognizing aspects of logic far from trivial.

Dodgson's confusing paradox about three barbers, based on ambiguity about the usage of 'if, then', aroused widespread controversy. "What Achilles Said to the Tortoise", a paper in *Mind*, raised a deep question about the ultimate validity of all logic and mathematics. Such reasoning rests on the premises of formal systems, but can we be certain the premises are valid? To justify them one must make more fundamental assumptions, and so on into the abyss of an infinite regress.

Dodgson was constantly printing leaflets and pamphlets about recreational mathematics and word games. Cohen surveys them all: original board games, games with cards, words and numbers, cipher methods, a mnemonic system for memorizing dates and large numbers, divisibility rules, improved systems of voting, and rules for croquet, for tennis tournaments, for playing billiards on a circular table and for rapidly determining a given date's day of the week.

Geologists and life scientists may be surprised to learn that Dodgson accepted an ancient Earth and evolution as God's way of creating life. He inclined toward the 'jump' theory proposed by the British zoologist St George Mivart, a Roman Catholic who was excommunicated for his heretical opinions. Cohen quotes from Dodgson's diary: "read the whole of Mivart's *Genesis of Species*, a most interesting and satisfactory book, showing as it does, the insufficiency of 'Natural Selection' alone to account for the universe, and its perfect compatibility with the creative and guiding power of God."

In many other ways Dodgson was theologically more liberal than most contemporary Anglicans. He totally rejected, for example, the doctrine of eternal punishment for the wicked. "If anyone urges 'then, to be consistent, you ought to grant the possibility that the Devil himself might repent and be forgiven'", he wrote in a letter, "I reply 'and I do grant it!'".

Literature on Dodgson continues to proliferate. Cohen's book is only the latest and most definitive of many biographies. There is even a biography by Anne Clark of Alice Liddell. Carroll societies flourish in England, the United States and Japan, each issuing a periodical. Unexpurgated

editions of Dodgson's diary, in many volumes, are now being edited by Edward Wakeling.

The Romans used a white stone or piece of chalk to mark on a calendar a day that gave special pleasure. The Latin phrase "a white stone day" was equivalent to what now is called a red letter day. Whenever an event was specially memorable to Dodgson, as on 25 April 1856, when he first met Alice (she was then almost four), he would write in his diary: "I mark this day with a white stone". Throughout the world Carrollians will be marking with white stones the day they begin reading Cohen's marvellous book. □

Martin Gardner, author of The Annotated Alice, More Annotated Alice, The Annotated Snark and other books about science, mathematics, literature and philosophy, is at 3001 Chestnut Road, Hendersonville, North Carolina 28792, USA.

■ Macmillan has just published the first full-colour edition of Lewis Carroll's *Alice's Adventures in Wonderland*. The original edition, first published by Macmillan in 1866, contained 42 black-and-white line drawings by the respected *Punch* illustrator Sir John Tenniel. In 1911, with Tenniel's approval, Macmillan commissioned the artist Harry Theaker to create eight colour plates from the original drawings. For this new hardback edition, Diz Wallis has coloured the remaining illustrations in the style of Theaker. Price is £12.99.

From five stars down

Keith Devlin

Five Equations that Changed the World: The Power and Poetry of Mathematics. By Michael Guillen. Little, Brown: 1995. Pp. 277. \$22.45, £18.99.

Five Golden Rules: Great Theories of Twentieth-Century Mathematics and Why They Matter. By John L. Casti. Wiley: 1995. Pp. 225. \$24.95, £16.99.

Nature's Numbers: Discovering Order and Pattern in the Universe. By Ian Stewart. BasicBooks/Weidenfeld and Nicolson: 1995. Pp. 164. \$20, £9.99.

Pythagoras' Trousers: God, Physics, and the Gender Wars. By Margaret Wertheim. Times: 1995. Pp. 279. \$23.

Conversations on Mind, Matter, and Mathematics. By Jean-Pierre Changeux and Alain Connes. Princeton University Press: 1995. Pp. 260. \$24.95, £19.95.

FOR those of us old enough to remember the 1960s, the phrase 'Five Easy Pieces' signifies a Jack Nicholson movie of that title. The same phrase ought by rights to apply to these five recently published

mathematics books, all of which are aimed at the general reader. But it does not. Both the quality and the accuracy of the writing and the level of sophistication and commitment required of the reader vary enormously across the five.

Part of the difficulty in writing a mathematics or science book for a general audience is that no-one knows exactly who the so-called 'general reader' is. Undeterred, both publishers and authors have been chasing this elusive character ever since Stephen Hawking hit the royalties jackpot with *A Brief History of Time*. (In fairness to publishers, valiant attempts were made to convey mathematics to the general reader before Hawking. For example, Penguin published Ian Stewart's *The Problems of Mathematics* before the 'Hawking effect' came into play and a revised edition is due to be republished by Oxford University Press in April as *From Here to Infinity*.)

With no-one quite sure who is reading their books, authors and publishers have to rely on their best guess at how to pitch their wares. Michael Guillen clearly assumes his readers are near-morons. *Five Equations that Changed the World* is without doubt the worst book on mathematics I have ever encountered. The breathless prose on the dust-jacket informs us that this book "tells the amazing stories of the people and the discoveries that led to the five most powerful and important scientific achievements in human history". Guillen, we are told, "reveals in accessible yet unpatronising language the secret world of mathematics".

The reader encounters Guillen's unpatronising language very early on. On page 10 we find the following passage about the young Isaac Newton: "As he walked along slowly, his brain raced toward a solution. Suddenly, however, he felt a sharp pain in his gut; his thoughts came to a screeching halt. As his mind's eye refocused, young Newton came out of his daydream and beheld his worst nightmare: Arthur Storer, the sneering, taunting school bully, had just kicked him in the stomach." This particular episode goes on in this way for four agonizing pages, doubtless causing Newton to turn in his grave and all self-respecting science writers to throw up their hands — and Guillen's book — in despair.

The blurb on the cover tells us that Guillen teaches mathematics at Harvard University, although a quick glance at my reference sources indicates that he is not listed among the regular mathematics faculty there. He is, I believe, a science correspondent for one of the major US television networks, which might explain his appeal to a publisher. My own experience with television and radio is that the people who work for them are generally extremely bright and creative, and I have no reason to suppose this is not true of Guillen. What a waste then to direct that

talent, and his influential position, to packaging mathematics and science into the genre of a badly written, trashy novel. The book is littered with factual inaccuracies as well. One wonders how the publisher resisted the temptation to put an embossed Fabio-like Galileo on the front cover.

For an example of first-rate popular mathematics writing, the discerning bookstore customer should quickly pass over Guillen's book and opt instead for Ian Stewart's *Nature's Numbers*. In this short book — it is only 150 pages long — Stewart achieves what other popular mathematics writers merely strive for: an accurate, informative portrayal of contemporary mathematics without a single equation in sight. Stewart's book is one of the Science Masters series assembled by the flamboyant and controversial literary agent John Brockman, famous for his 'big money deals'. The books in this series prove that good-quality accurate science writing can be both accessible to a general audience and marketed with a keen eye to making a tidy profit for the author and the publisher (and in this case the agent as well). If you are a mathematician or a scientist and someone you know wants to know what mathematics really is, buy them a copy of *Nature's Numbers*. Stewart does not try to explain very much actual mathematics. Instead, he describes what mathematics is, how it is done and how it relates to ourselves and the world we live in.

An equally smooth prose ride is offered by Margaret Wertheim in *Pythagoras' Trousers*. Wertheim is a science writer with bachelor's degrees in physics and in mathematics and computing. Her writing is among the best I have ever had the pleasure to read. I was captivated by the first few paragraphs of her introduction, and I urge Guillen to read this book to see a true example of "accessible yet unpatronising language".

Wertheim takes the reader through the development of physics, starting with the mathematics and astronomy of ancient Greece, and moving through Bacon, Copernicus, Galileo and all the rest of the seventeenth-century crowd, up to the present day. As a well-researched, historical account I found it fascinating, and I was intrigued by her theme that today's male domination of physics has its roots in the historically close connections between physics and Christianity. According to Wertheim, physics has always seen itself as a sort of male-dominated priesthood. Physics, says the book's dust-jacket, is "the Catholic church of science", with all that it entails.

I must admit that, although physics is still a heavily male-dominated field, I am not convinced that the explanation is to be found in Wertheim's religious-cult theory — it is certainly not based on the physi-

cists I know. I definitely do not agree with her that the recent rash of physics books with the word 'God' in the title shows that physicists view themselves as a priesthood engaged in a religious enterprise. I suspect the answer has more to do with the Hawking effect: having 'God' in the title helps to sell books, and that's all there is to it.

I was also disturbed by Wertheim's apparent confusion of physics with technology, and the conclusions she draws as a result: that society should direct physicists to concentrate their efforts on "helping humanity" in the most obvious and immediate ways. On this point, I feel Wertheim misunderstands the nature of science, and her claim that much of today's physics research will be of no benefit to mankind seems odd from one who has so obviously studied the history of science.

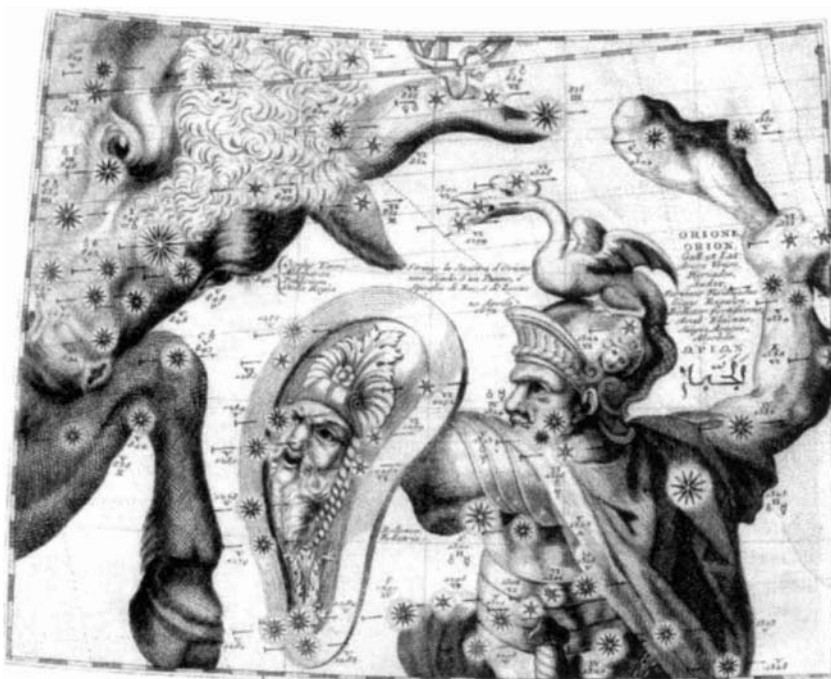
But whether or not you agree with Wertheim's thesis, her book is a delight to read, and highly informative. So too is *Conversations on Mind, Matter, and Mathematics*, translated from the original French by M. B. DeBevoise. The book is an edited transcript of a series of conversations between a leading neurobiologist, Jean-Pierre Changeux, and a Fields Medal-winning mathematician, Alain Connes. The book's title sums up what the two talked about: the nature of mathematics and how the human brain can engage in mathematical thought about abstract entities.

As an insight into the views of two of the world's leading scientists, the book is a success. As a conversation, it is in some ways a failure. What struck me throughout was that, for all Connes's attempts at an explanation, Changeux never does understand the real nature of mathematics, and as a result much of their conversation is at cross-purposes.

And so to the fifth and final book, John Casti's *Five Golden Rules*. In many ways this is not really aimed at the 'general reader'. It is hard going from the beginning, although well worth the investment for anyone who really wants to learn about some of the contemporary mathematics that affects our everyday lives. The book's high-profile marketing probably results from Casti's earlier bestseller *Paradigms Lost* (also not the easiest of reads) and the fact that he is a cracking good writer. The five themes chosen by Casti are game theory, topological fixed-point theory, singularity theory, computation theory and optimization theory. He makes no attempt to entice you in. He assumes you want to know. For those who do, the reward is substantial. □

Keith Devlin is at Saint Mary's College of California, PO Box 3517, Moraga, California 94575-3517, USA.

Stars in their eyes



STARMAN — the constellations of Orion and Taurus mapped by Venetian Vincenzo Coronelli around 1701. Coronelli produced unique hand-painted celestial globes that were scientifically impeccable but mostly intended to be artistic: the positions of the stars are almost lost in the intricate engraving of the figures of the constellations. It is one of many illustrations in Peter Whitfield's *The Mapping of the Heavens*, which charts mankind's conceptualization of the heavens through cartography from the stone age to the space age. Pomegranate/British Library Press, \$29.95, £20.

Crossing cultural thresholds

Lucy Bending

The History of Pain. By Roselyne Rey. Harvard University Press: 1995. Pp. 394. \$39.95, £25.50.

PAIN is not simply a brute given, a constant to be interpreted always in the same way. It is pulled sharply this way and that by culture and by expectation. Although bound to the body, the experience of pain is hedged about by frameworks of explanation, which in turn affect the way in which the pain is felt. Nietzsche, calling his pain "dog", made sense of his physical suffering by fitting it into just such a framework: "It is just as faithful, just as obtrusive and shameless, just as entertaining, just as clever as any other dog — and I can scold it and vent my bad mood on it, as others do with their dogs, servants, and wives".

Roselyne Rey in this serious and scholarly book, aimed at the historian of medicine rather than the medical student, recognizes the role of culture in shaping the perception of bodily suffering. She examines medical ideas about bodily pain

from Graeco-Roman times until the 1950s, delineating and questioning the ways in which physicians have attempted to get to grips both with the mechanics of physical suffering and with formulating solutions to the problems it poses. This history does not suggest a steady accumulation of knowledge, but rather provides a survey of continuities and discontinuities, as dissection fell in and out of favour, or as classical anatomical texts were reformed in the Renaissance, reopening long-forgotten paths of investigation. Methods of research may have progressed, as staining techniques and microscopy made neurological investigation possible, or the introduction of gunpowder into warfare forced advances in understanding, but frameworks of reference may also subsist beyond their time, serving to hold back medical advance. As Rabbi Julia Neuberger suggested in the 1980s, old ideas of sin and its connectedness to pain are fed back into modern ideas of natural childbirth, tearing themselves away from their original religious context, and suggesting needlessly that it

is good to suffer for suffering's sake.

It is in the perception of such chronological ambiguities that Rey's greatest strengths lie. Just as individual perception of pain is dependent on culture, so too were medical investigators of pain constrained by their cultural and scientific milieu. Ronald Pearsall, writing in 1975, singularly failed to acknowledge cultural forces at work, recognizing only "laziness and apathy" in the doctors of the late 1840s who refused to change their practice when faced with the possibility of anaesthetic surgery. In his eyes, their failure to use chloroform resulted "in agony to patients that was totally unnecessary".

Rey avoids such crassness. Of course some doctors were held back by apathy, but, as Rey points out, things that seem self-evident after the passage of years are by no means so simple at the time. Physicians of previous centuries — or indeed decades — did not see what seems obvious to us now. This was not through obstinacy or stupidity, but because of their necessary engagement with the critical thought of the time. Seeing laughing-gas in the same analgesic terms as post-operative opium, they were unable to recognize the potential of the gas as a surgical anaesthetic. Those who first used chloroform in surgery overcame this conceptual difficulty, but in so doing took a "political decision to take a chance on surgery without pain and, by so doing, took upon themselves the risk of the odd fatal accident". The 25,000 deaths attributed to anaesthetics in England and Wales in the century after 1846 make this point loudly and clearly.

It is clear, too, from these figures — as indeed from the heated debate raging in the pages of the *Lancet* in the mid-nineteenth century — that the decision to use anaesthetics in surgery was by no means a foregone or 'natural' conclusion. Older frames of reference persisted as doctors on the battlefield refrained from using chloroform — not just because using chloroform would have required another doctor to act as anaesthetist, effectively halving the number of surgeons in the field, but also because to be brave in the face of pain was manly, and to desire the removal of such suffering was a sign of weakness. Such a desire to suffer, Rey makes clear, is not confined to any one period. Motives change and times change, yet both the flagellants of the Middle Ages as well as the French 'dolorists' between the two world wars "saw in pain a type of catharsis, a means of purification from non-essentials, incidentals and falsehoods". The analgesic function of medicine, faced with such an attitude, fades into insignificance.

If Rey reaches one conclusion in her clear overview, it is that the way in which

we construct explanations for pain affects the way in which we perceive it, possibly with adverse effects on the sufferer. Although the book is staunchly academic, Rey manages to tell the human story of pain, recognizing that while pain is a fitting subject for academic research, it is also felt, both physically and emotionally, by the individual sufferer. As the anniversary of her death approaches, this splendid book seems a fitting memorial to the work of a woman who died from breast cancer at the age of 44. □

Lucy Bending is at St Hugh's College, Oxford OX2 6LE, UK.

Generating diversity

Nel C. Moore

T Cell Receptors. Edited by John I. Bell, Michael J. Owen and Elizabeth Simpson. Oxford University Press: 1995. Pp. 483. £60, \$53 (hbk); £29.50, \$107 (pbk).

OUR understanding of the structure and function of T-cell antigen receptors (TCR) has come a long way since the discovery of these proteins just over a decade ago. Much of this progress is reviewed in this comprehensive book, which should appeal to a wide audience.

TCR α - and β -chains, the variable (V) regions of which are responsible for antigen recognition, are encoded by discontinuous gene segments that are rearranged specifically in T cells during development. Detailed information about the extent and variability of the germline repertoire of V gene segments is therefore essential for an understanding of the potential diversity of T-cell specificity. The several chapters dedicated to this topic and the processes governing the rearrangement of these gene segments clearly emphasize the complexity of these events. In this respect, the inclusion of a table and nomenclature-listing of TCR α - and β -chain V regions, as well as a plate section comparing various human and mouse V-region peptide sequences, are very helpful.

Unlike B cells, T cells can recognize only antigens presented by major histocompatibility complex (MHC) molecules on the surface of antigen-presenting cells. This so-called MHC-restriction, which results from selection processes during T-cell development in the thymus, represents the only specific interaction between T cells and antigen-presenting cells. So it is appropriate that a detailed structural and functional analysis of both TCR-MHC-peptide and TCR-MHC-superantigen complexes occupies several

excellent chapters. Similarly, the importance of these complexes in the development of the TCR repertoire in the thymus is also discussed.

Many other TCR-related subjects are covered. There is an elegant description of T cells expressing $\gamma\delta$ TCRs rather than $\alpha\beta$ TCRs, although the precise immunological function of $\gamma\delta$ T cells is not yet fully known. And the impressive progress in understanding of intracellular signalling processes triggered by the engagement of antigen with TCR is also described in detail.

Inevitably in a multi-authored book such as this there is some repetition as well as a lack of coverage of several recent advances. Nevertheless, the volume is a solid work of reference that will be particularly valuable to those either starting out in or wanting to refresh their knowledge of T-cell immunology. □

Nel C. Moore is in the Department of Anatomy, Medical School, University of Birmingham, Birmingham B15 2TT, UK.

New in paperback

Power Unseen: How Microbes Rule The World by Bernard Dixon. W.H. Freeman/Spektrum, £9.99. "...covers the entire microbial gamut in 75 short essays ... informative... entertaining." Robert Desowitz, *Nature* **368**, 820 (1994).

Henderson's Dictionary of Biological Terms edited by Eleanor Lawrence. Longman, £7.99 (pbk). Now in its eleventh edition, this handy dictionary has been thoroughly updated and revised to include more than 23,000 references.

The Elements by John Emsley. Oxford University Press, £10.99. The second edition of this useful reference book in the Oxford Chemistry Guides series now appears in a 'pocket' format with a plastic 'flexicover'.

Bicycling to Utopia: Essays on Science and Technology from the Royal Institution edited by Peter Day and Richard Catlow. Oxford University Press, £9.99. A collection of selected "Friday Evening Discourses" given at the Royal Institution in London. The essays include Steve Jones on the future of human evolution, Iwao Fujimasa on artificial hearts, Malcolm Wilkins on plant intelligence and D. T. Clark on the materials used in Formula 1 racing.

The Biochemists' Songbook by Harold Baum. Taylor and Francis, £6.99. The second edition of a classic work in which "important topics of biochemistry have been rendered... into light-hearted songs and to appropriate popular tunes... intended for communal singing, ideally with a blood alcohol level of 35 mg per cent" — at least according to Thomas Scott writing in *Nature* **296**, 513 (1982).

Requirement for *Xist* in X chromosome inactivation

Graeme D. Penny, Graham F. Kay*, Steven A. Sheardown, Sohaila Rastan* & Neil Brockdorff†

Section of Comparative Biology, MRC Clinical Sciences Centre, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

The *Xist* gene has been proposed as a candidate for the X inactivation centre, the master regulatory switch locus that controls X chromosome inactivation. So far this hypothesis has been supported solely by indirect evidence. Here we describe gene targeting of *Xist*, and provide evidence for its absolute requirement in the process of X chromosome inactivation.

DOSAGE compensation of X-linked genes in mammals is achieved by the transcriptional silencing of a single X chromosome early in the development of the female embryo. This process, termed X chromosome inactivation (X inactivation), is normally random, with an equal probability that either X chromosome will be inactivated in a given cell¹. Recent progress towards understanding X inactivation has been made through attempts to identify the X inactivation centre (Xic), a *cis*-acting locus on the X chromosome which is the master regulatory switch controlling this process².

Initiation, the first step in X inactivation, involves the cellular determination of how many (if any) and which X chromosomes to inactivate. Experimental observations on X; autosome translocations and X chromosome deletions suggest that the Xic is required for the initiation of X inactivation in early embryogenesis, there being an absolute requirement for at least two Xics to be present^{3,4}. Further evidence supporting a role for the Xic in the initiation of X inactivation comes from studies of alleles at the X controlling element (*Xce*) locus, which maps to the Xic region on the mouse X chromosome^{5–7}. *Xce* homozygotes undergo normal, random X inactivation, whereas heterozygotes show primary non-random X inactivation⁸, as there is a greater probability that an X chromosome bearing a strong *Xce* allele will remain active. Thus the Xic is implicated both in the decision as to how many X chromosomes are inactivated (counting), and also which X chromosomes are inactivated (choice). Studies on X-chromosome rearrangements have also demonstrated that the Xic is required *in cis* for the spread of inactivation (propagation) to occur along the entire X chromosome^{9,10}. After the inactive state has been established, X inactivation patterns are stably maintained in all subsequent cell divisions. The inactive state is stably maintained after the loss of the entire minimal XIC region on the human X chromosome¹¹, which suggests that, although the Xic is required for initiation and propagation, it is not required for the maintenance of X inactivation.

A candidate for the Xic, termed *XIST* (X-inactive specific transcript), has been identified within the human minimum XIC interval^{12,13}. In contrast to all other X-linked genes, *XIST* is expressed exclusively from the inactive X chromosome. The equivalent mouse gene (*Xist*) was cloned through homology, localized to the mouse Xic interval, and shown to be expressed exclusively from the inactive X chromosome^{14,15}. The *XIST/Xist* gene encodes a large RNA transcript which lacks protein coding potential and is localized in the nucleus, apparently in direct association with the inactive X chromosome^{16,17}. This finding has

led to the proposal that the transcript acts *in cis* to propagate X inactivation. Alternative models remain tenable, for example that the *XIST/Xist* locus represents a nucleation centre for heterochromatin induction that is initiated by transcription of the locus¹⁶. Studies of expression of the mouse gene in early embryogenesis^{18,19}, and in XX embryonic stem (ES) cells¹⁸, which undergo X inactivation on differentiation *in vitro*³, demonstrate that *Xist* expression precedes overt X inactivation, consistent with it having a role in the initiation process. *Xist* is not expressed in male tissues other than male germ cells just before meiosis^{18,20,21}. This observation has led to the suggestion that *Xist* could also be involved in the transcriptional silencing of the XY bivalent in the sex vesicle.

We have tested directly the hypothesis that *Xist* is required for X inactivation, by introducing a targeted deletion into a single *Xist* allele in an XX embryonic stem (ES) cell line, and have investigated the effect of this mutation on X inactivation following differentiation of mutant cells *in vitro* and *in vivo*. At the outset we predicted three possible outcomes: (1) that mutant cells would fail to register the presence of two Xics, and would entirely fail to undergo X inactivation; (2) that the mutation would disrupt X inactivation such that the targeted X chromosome would fail to X inactivate, although the normal X chromosome would undergo X inactivation; and (3) that the mutation would have no effect on X inactivation at all. Our results show that the second prediction is correct, and that X inactivation fails to occur *in cis* on the X chromosome bearing the deleted *Xist* allele.

XX ES cells: an *in vitro* model

We targeted *Xist* in XX ES cells, as this allows analysis of the initiation of X inactivation by *in vitro* differentiation of mutant ES cells. We derived an XX ES cell line (PGK12.1) using delayed blastocysts from a 129 female crossed to a (PGK × 129)F₁ male. The PGK12.1 cell line is thus heterozygous for a PGK and a 129 derived X chromosome. Because the X chromosome in the C3H/He-Pgk-1^a/Ws (PGK) mouse is derived from a wild mouse strain²², it is relatively easy to detect polymorphisms for X-linked markers to assist in subsequent analysis.

We have previously shown that the PGK12.1 ES line expresses both *Xist* alleles and undergoes random X inactivation upon differentiation *in vitro*²³. However, analysis of the methylation status of several CpG sites at the 5' end of the *Xist* gene indicated that undifferentiated XX ES cells are methylated at approximately 50% of the level seen in XY ES cells, which show near-complete methylation. This observation led us to hypothesize that only a single *Xist* allele is methylated in an individual XX ES cell, and that this may reflect predetermination with respect to *Xist* expression and X inactivation²³. As predetermination of individual ES cells would have important implications in terms of interpreting the outcome of *Xist* gene targeting, we have investi-

* Present addresses: Queensland Institute of Medical Research, The Bancroft Centre, 300 Herston Road, Brisbane, Q4029, Australia (G.F.K.); and The Medical Research Council, 20 Park Crescent, London W1N 4AL, UK (S.R.).

† To whom correspondence should be addressed.

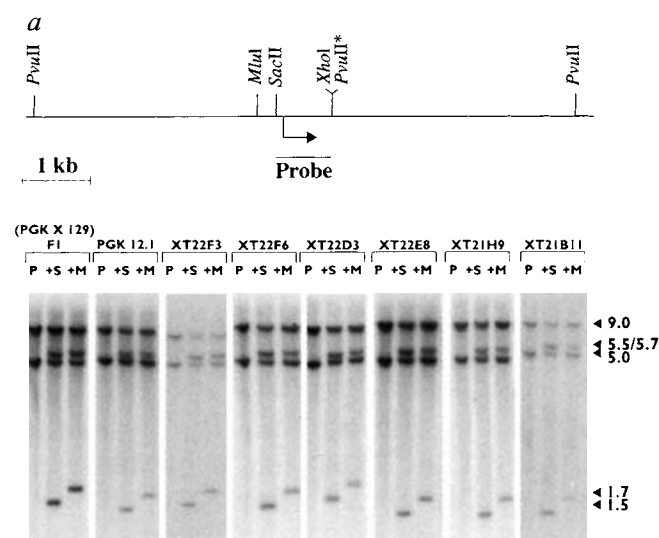
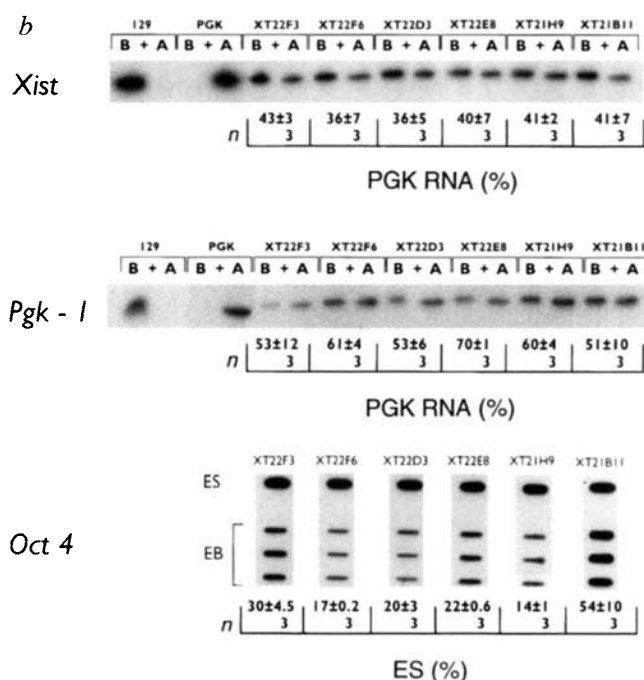


FIG. 1 Clonal PGK 12.1 XX ES cells are not predetermined with respect to *Xist* expression and X inactivation. **a**, Southern analysis demonstrating methylation of both *Xist* alleles in six clonal ES cell lines (XT22FG3, XT22F6, XT22D3, XT22E8, XT21H9 and XT21B11), derived from the PGK12.1 ES cell line. A polymorphic *PvuII* site, indicated on the map by an asterisk, was used to distinguish the PGK (5 kb), and 129 (9 kb) alleles. The transcription start site is indicated with an arrow. After digestion with *PvuII* (P), samples were digested with the methylation-sensitive restriction enzymes *SacII* (+S) or *MluI* (+M). The results indicate approximately 50% methylation of both alleles in all six clonal cell lines. Controls were (PGK × 129)_{F1} female mouse and non-clonal PGK12.1 ES cells. The probe used was the *SacII*/*XhoI* fragment indicated. Band sizes are marked in kilobases. **b**, The relative RNA levels of *Xist* and *Pgk-1* alleles (A, PGK; B, 129) in differentiated clonal PGK12.1 lines was determined by quantitative SNUPE analysis. Quantification of SNUPE assays is expressed for the PGK allele as the percentage of total $\pm$ s.e.m. The results indicate random *Xist* expression and X inactivation with a pronounced *Xce* effect. Controls are adult tissue RNA (129 or PGK). Three independent cultures were assayed for each clonal cell line. Differentiation was monitored by slot-blot analysis of *Oct4* RNA levels. Quantification is expressed for differentiated cells (EB) as percentage of ES cell level $\pm$ s.e.m. Downregulation of *Oct4* was less pronounced with XT21B11 cells than with the other lines. **METHODS.** ES cell line PGK12.1 was derived essentially as described previously³⁷, using delayed blastocysts from a 129^{ola/hsd} × (C3H/He – Pgk – 1^a/Ws × 129^{ola/hsd})_{F1} backcross. ES cells were cultured in ES cell medium; DMEM with 4.5 g l⁻¹ glucose, no sodium pyruvate, 2 mM glutamine, 1 × non-essential amino acids (GIBCO), 0.05 mg ml⁻¹ streptomycin, 500 i.u. l⁻¹ penicillin, 0.1 mM 2-mercaptoethanol, 20% fetal calf

gated this further by analysing six clonal PGK12.1 ES cell lines. If individual ES cells are predetermined, in clonal lines we would expect to see methylation of only one allele and non-random *Xist* expression and X inactivation following differentiation. Methylation analysis of *Xist* in the clonal cell lines is illustrated in Fig. 1a; both alleles show approximately 50% methylation levels. Expression of *Xist* and *Pgk-1* in differentiated cells was analysed using the quantitative allele-specific single nucleotide primer extension assay (SNUPE)^{24–26} (Fig. 1b). In this experiment the extent of differentiation was assessed by monitoring downregulation of *Oct4* RNA, an early marker of ES cell differentiation³⁷ (Fig. 1b). In each clone we observed expression of the 129 strain (B) and PGK strain (A) alleles of *Xist* and *Pgk-1*. Quantification of these results demonstrated lower levels of PGK strain *Xist* RNA and higher levels of PGK strain *Pgk-1* RNA, consistent with the PGK X chromosome carrying a stronger *Xce* allele (*Xce*^c) than the 129 X chromosome (*Xce*^a). This result accords with previous observa-



serum, and 10³ ug ml⁻¹ recombinant LIF (GIBCO), at 37 °C, 5% CO₂, 100% humidity. Clonal PGK12.1 lines were isolated by geneticin (400 μg ml⁻¹) selection following transfection with *Pgk-neo*. Methylation analysis was as described previously²³. Cells were differentiated using the embryoid body outgrowth method³⁷. Cells were first plated in EB medium: DMEM with 4.5 g l⁻¹ glucose, no sodium pyruvate, 2 mM glutamine, 0.05 mg ml⁻¹ streptomycin, 500 i.u. l⁻¹ penicillin, 0.1 mM 2-mercaptoethanol, 10% fetal calf serum, at 37 °C, 5% CO₂, 100% humidity. After 3 days, cells were transferred to suspension culture for 3–6 days, and then returned to monolayer culture for a further 3 days before collection. RNA was prepared using RNAzol reagent (Biogenesis). DNA isolation, Southern and northern slot-blot analysis were performed using standard methods. Slot blots were loaded with 5 μg RNA, and probed with an *Oct4* cDNA probe corresponding to bases 491–953 of the published sequence³⁸. Quantitative SNUPE analysis was performed as described previously²⁴. *Pgk-1* SNUPE analysis used primers and conditions described previously²⁴. For *Xist*, RT-PCR was as described previously¹⁸. SNUPE, 1 cycle of 94 °C 1', 50 °C 2' and 72 °C 1', was performed using primer Xist-9974; GGTTCTCTCCAGAAGCTAGGA, which detects a previously described polymorphism²⁰ (129, A; PGK, G). Quantification of SNUPE and *Oct4* RNA slot blots was performed on a Molecular Dynamics phosphorimager. All SNUPE results were normalized using RNA samples with known allelic levels²⁴.

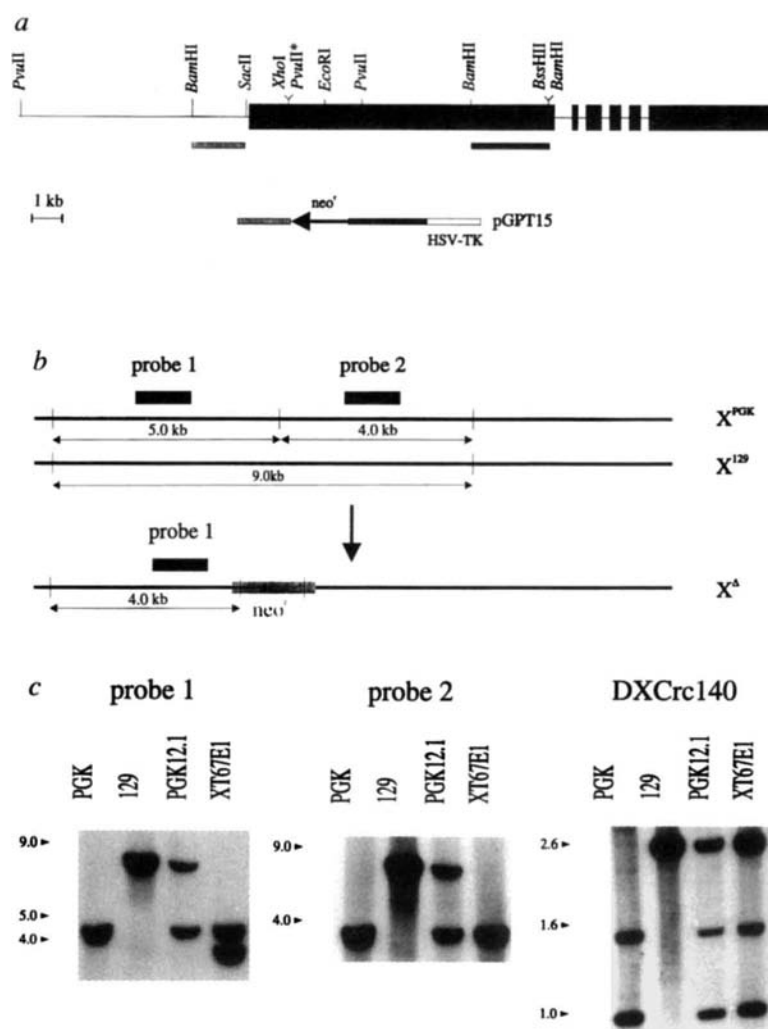
tions of an *Xce* effect in similar XX ES cell lines²⁶. These results demonstrate that clonal PGK12.1 cells are not pre-determined with respect to *Xist* expression or X inactivation.

Targeted mutagenesis

The gene-targeting strategy adopted is illustrated in Fig. 2. Because *Xist* RNA does not encode a protein, we created a large deletion of the *Xist* gene to maximize the probability of creating a null allele. The targeting construct, pGPT15, was assembled using 5' and 3' homologous fragments separated by approximately 7 kilobases (kb) of sequence from exon 1 of the *Xist* gene and including 36 base pairs (bp) of *Xist* minimal promoter sequence. The construct (Fig. 2a) includes a *Pgk-neo* gene for positive selection of stable integrants, and the herpes simplex virus thymidine kinase gene (HSV-TK) for selection against non-homologous events²⁸. We screened 2,588 stably transfected colonies by Southern analysis, and a single homologous integrant,

FIG. 2 Targeted deletion of the 129 strain *Xist* allele in PGK12.1 ES cells. **a**, A partial restriction map of the *Xist* locus indicating the gene structure and regions of homology used to construct the targeting vector pGPT15. **b**, A *PvuII* polymorphism was used to distinguish *Xist* alleles in PGK12.1 cells. The size of expected restriction fragments detected with probes 1 and 2 in the parent and targeted 129 loci are shown. **c**, Southern blot analysis of *PvuII*-digested control female tissue DNA (PGK and 129 strain), and PGK12.1 and XT67E1 ES cell line DNA, using probes 1 and 2, shows that the cell line XT67E1 carries a targeted deletion of the 129 strain *Xist* locus (sizes in kilobases). Probe 1 detects a 4-kb band, as predicted for homologous integration, and the 5-kb PGK-derived band indicating that the 129 allele has been targeted. Probe 2, which lies within the deleted region, only detects the 5-kb PGK allele. A *PvuII* polymorphism at the *DXCrc140* locus³⁹ was used to ascertain that XT67E1 cells had not eliminated the PGK strain X chromosome.

METHODS. The 5' (*Bam*H1/*Sac*II) and 3' (*Bam*H1/*Bam*H1) homologous fragments used in the pGPT15 construct were isolated from genomic clones from a 129 library in lambda DASH (Stratagene). PGK12.1 ES cells (1×10^7 – 2×10^7) were electroporated with 20 μ g of linearized pGPT15 (500 V/25 μ F in a Biorad Gene Pulser), and selection was then performed over 7 days in ES cell medium containing 400 μ g ml⁻¹ geneticin and 2 μ M gancyclovir. Individual colonies were picked onto 96-well plates. Replicate plates were prepared for isolation of DNA and Southern analysis as described previously⁴⁰. Southern blots were probed with radiolabelled *Xist* probes as indicated, and also with the pEM140A probe from the *DXCrc140* locus³⁹.



XT67E1, was identified. Southern analysis, using a *PvuII* polymorphism to distinguish the 129 and PGK *Xist* alleles (Fig. 2b), demonstrated that the targeting event occurred on the 129-derived X chromosome (Fig. 2c). Further analysis with a 3' probe, and also with a range of restriction enzymes, was used to confirm this result and demonstrate correct homologous integration (not shown).

X inactivation in vitro

Xist expression was analysed in differentiated derivatives of XT67E1 cells, using a *Hind*III polymorphism that allowed transcripts from the two alleles to be distinguished¹⁸. As expected, differentiated XT67E1 cells only expressed *Xist* from the non-targeted allele (Fig. 3a), confirming that the deletion of the 129 *Xist* locus results in a null allele. This finding also suggests that the counting function of the Xic is not disrupted in targeted cells, that is, XT67E1 cells retain the ability to register the presence of two Xics. However, disruption of the counting mechanism cannot be precluded from this result alone as it is not known formally that *Xist* expression is an absolute indicator of the decision to undergo X inactivation. For this reason, differentiated cells were analysed for the presence of an inactive X chromosome by using a cytogenetic late-replication assay^{29–31}. Metaphase spreads were prepared from control female BALB/c cultured thymocytes, and from PGK12.1 and XT67E1 ES cells and their differentiated derivatives, and were assayed for the presence of a late-replicating X chromosome in a double-blind screen (Fig. 3b). Late-replicating

X chromosomes were not found in undifferentiated ES cells, consistent with there being two active X chromosomes, but were observed in both PGK12.1 and XT67E1 differentiated cells. This result confirms that the *Xist* deletion does not result in loss of the counting function of the Xic. We did, however, observe a slightly reduced frequency of late-replicating X chromosomes in differentiated XT67E1 (29%) compared with PGK12.1 cells (44%) (see discussion).

Nonrandom X inactivation in vitro

To determine whether targeted cells undergo normal random X inactivation, we used the SNUPE assay to assess the relative levels of X-linked gene products in PGK12.1 and XT67E1 ES cells and their differentiated derivatives. In addition to the precharacterized polymorphism used to assess the relative expression level of the two alleles at the *Pgk-1* locus (Fig. 1b), we also characterized new polymorphisms to assess the expression levels of the PGK and 129 alleles of the *Rps4* gene and also the *Smcx* gene, which escapes X inactivation in adult mouse tissues^{32,33}. A polymorphism in the *Zfx* gene was used to quantify genomic DNA.

A summary of results from the SNUPE analysis is shown in Fig. 4a. Analysis of all three genes in PGK12.1 and XT67E1 undifferentiated ES cells shows that both alleles are expressed equally, as would be expected for cells with two active X chromosomes. In tissues from control (129 $\times$ PGK) F₁ females and differentiated parent PGK12.1 cells (EB), a significantly higher level of PGK strain RNA was observed, consistent with the

expected *Xce* effect. In marked contrast, a significantly lower level of transcripts from the PGK strain allele was observed for all three genes in differentiated targeted XT67E1 cells. Quantitative slot-blot analysis of *Oct4* RNA (Fig. 4b) indicates extensive, and equivalent, differentiation of PGK12.1 and XT67E1 cultures ($13.7 \pm 3\%$, $n = 4$, and $16.8 \pm 3.8\%$, $n = 5$, of mean ES cell value, respectively). The *Zfx* DNA controls indicate that the skewed allelic RNA levels in XT67E1 cells are not due to specific elimination of the PGK X chromosome. In adult tissues the *Xce* effect is considerably reduced for the *Smcx* gene, consistent with the fact that this gene escapes X inactivation, whereas in differentiated PGK12.1 cells the full *Xce* effect is apparent. This suggests that the *Smcx* allele on the inactive X chromosome undergoes near-complete X inactivation on differentiation of ES cells. This is supported by recent observations demonstrating that the *Smcx* gene only partly escapes X inactivation immediately after the onset of X inactivation *in vivo* (S.A.S., manuscript in preparation).

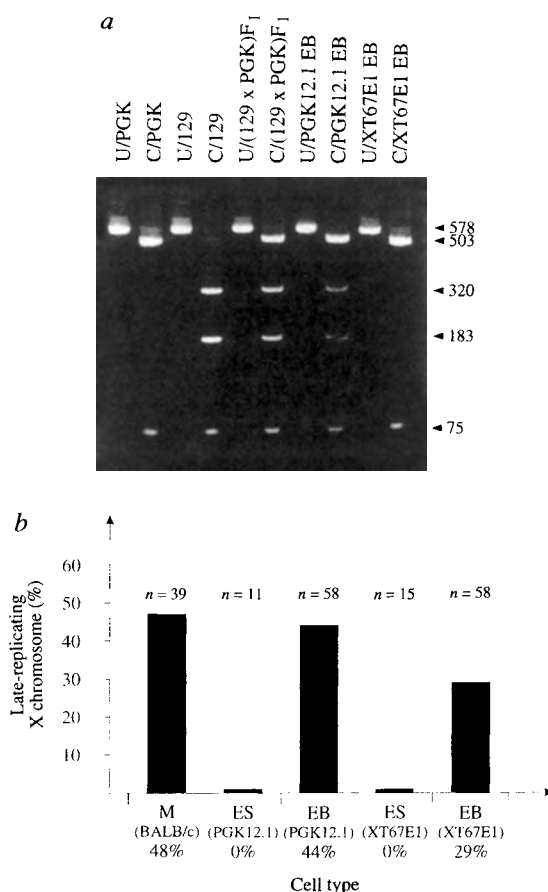
The allelic ratios for X-linked RNAs in XT67E1 cells is therefore skewed in the opposite direction to the *Xce* effect expected and observed in control PGK12.1 cells. To determine the basis for this, we performed SNUPE analysis of *Pgk-1* RNA following reverse transcription-polymerase chain reaction (RT-PCR) on individual cells (examples are shown in Fig. 4b). Both alleles of *Pgk-1* were detected, frequently at nearly equal levels, in all undifferentiated PGK12.1 and XT67E1 ES cells. Monoallelic expression, either of the PGK or the 129 strain allele, was observed in control cells from an adult female (PGK $\times$ 129) F_1 mouse (Fig. 4b, cells 7–12), and also from differentiated PGK12.1 cells (Fig. 4b, cells 15, 16 and 18), consistent with random X inactivation. In marked contrast, XT67E1 cells showed monoallelic expression of the 129 allele only (Fig. 4b, cells 19, 21, 22, and 24), indicating complete non-random X inactivation of the PGK derived X chromosome. A summary of

the single-cell analysis is shown in Fig. 4c. SNUPE assays were quantified, and cells showing $> 80\%$ levels of one allele were scored as having undergone X inactivation (see Fig. 4 Methods). Cells with nearly equal levels of both alleles ($50 \pm 10\%$) were scored as having two active X chromosomes. This analysis clearly demonstrates that differentiated PGK12.1 and control bone-marrow cells undergo random X inactivation, whereas XT67E1 cells undergo complete non-random X inactivation of the X chromosome bearing the non-targeted *Xist* allele. The proportion of cells with no inactive X chromosome was higher in the differentiated XT67E1 than the PGK12.1 cultures (see Discussion).

X inactivation *in vivo*

To analyse further the phenotype of XT67E1 cells, we assessed X inactivation *in vivo* by making aggregation chimaeras of parent PGK12.1 or targeted XT67E1 cells with eight-cell CD1 host embryos. RNA isolated from chimaeric embryos at 10.5–12.5 days post coitum (d.p.c.) was analysed using the SNUPE assay. Examples of the results obtained are shown in Fig. 5. Parent PGK12.1 cells exhibit random X inactivation in chimaeric embryos, as indicated by the expression of both alleles of *Xist* and *Pgk-1* in a PGK12.1 XY host chimaeric embryo. The high relative level of the 129 allele of *Pgk-1* and *Smcx* is due to the host embryo contribution (the CD1 and 129 X chromosome carry the same alleles). Random X inactivation of PGK12.1 cells, as indicated by expression of the PGK allele of *Xist* and *Pgk-1*, was also observed in two XX host chimaeric embryos (not shown). In contrast to this, examination of targeted XT67E1 cells demonstrates complete non-random X inactivation of the PGK strain X chromosome bearing the intact *Xist* locus. Thus, in XY host chimaeric embryos, only the PGK strain *Xist* allele is detectable owing to the absence of host embryo expression. In XX host chimaeric embryos, expression of the CD1 strain *Xist* allele from the host embryo is seen, as

FIG. 3 XT67E1 cells undergo initiation of X inactivation on differentiation *in vitro*. a, PGK12.1 and XT67E1 cells were analysed for *Xist* expression following differentiation. A *HindIII* polymorphism was used to distinguish 129 and PGK strain *Xist* alleles²⁰ (U, uncut; C, cut with *HindIII*). Band sizes are in base pairs. Controls were adult female tissue RNA from PGK and 129 strain, and a (129 $\times$ PGK) F_1 mouse. Differentiated (EB) PGK12.1 cells expressed both *Xist* alleles, and XT67E1 cells only the PGK strain allele, consistent with deletion of minimal promoter sequences from the 129 allele. b, Presence of a late-replicating inactive X chromosome in differentiated XT67E1 cells was demonstrated by a cytogenetic assay²⁹. Positive controls were differentiated PGK12.1 cells, and primary cultured thymocytes from a female BALB/c mouse. Negative controls were provided by undifferentiated PGK12.1 and XT67E1 ES cells. A double-blind experiment, scoring metaphase spreads with 40 chromosome karyotype for the presence or absence of a pale-staining (late-replicating) X chromosome was performed. No late-replicating X chromosome was detected in spreads from XT67E1 and PGK12.1 ES cells (both X chromosomes active), but in their differentiated derivatives (EB), and also control primary thymocytes from female BALB/c mice (M), late-replicating inactive X chromosomes were observed. The frequency of late-replicating chromosomes in differentiated PGK12.1 controls was consistent with that determined in previous studies^{28,30}, but was reduced in differentiated XT67E1 cells. METHODS. ES cells were differentiated as described in Fig. 1. The allele-specific *Xist* RT-PCR assay was performed as described previously¹⁸. The late-replication assay was a modification of that described previously²⁴. Cells were cultured with $200 \mu\text{g ml}^{-1}$ BrdU for 8 h and then 50 ng ml^{-1} colchicine for 1 h. BALB/c thymocytes were cultured in EB medium with $4 \mu\text{g ml}^{-1}$ concanavalin A for 3 days before addition of BrdU. Metaphase spreads were prepared by dropping acetic acid/methanol-fixed cells onto microscope slides. After ageing for several days, slides were stained with Acridine Orange in Gurr's buffer, pH 6.8, mounted and viewed under fluorescence on a Zeiss Axiophot. Slides were coded and then independently scored by two people.



well as the PGK strain *Xist* allele from donor cells. Analysis of the *Pgk-1* gene, however, demonstrates complete absence of expression of the PGK strain allele in XY and XX host chimaeric embryos, demonstrating that donor cells have all inactivated the PGK strain-derived X chromosome. The *Smcx* gene, which partly escapes X inactivation, is expressed from both alleles. The higher relative level of 129 alleles of *Pgk-1* and *Smcx* is again due to host embryo contribution. Identical results were obtained in three additional chimaeric embryos (not shown). The *in vivo* analysis

therefore confirms the *in vitro* findings, and clearly demonstrates that XT67E1 cells do not inactivate the 129 X chromosome bearing the deleted *Xist* allele.

Discussion

We have generated an XX ES cell line, PGK12.1, which carries genetically distinct PGK and 129 strain X chromosomes. We analysed clonal derivatives of this cell line and showed that both alleles of *Xist* and the *Pgk-1* gene are expressed at levels consistent

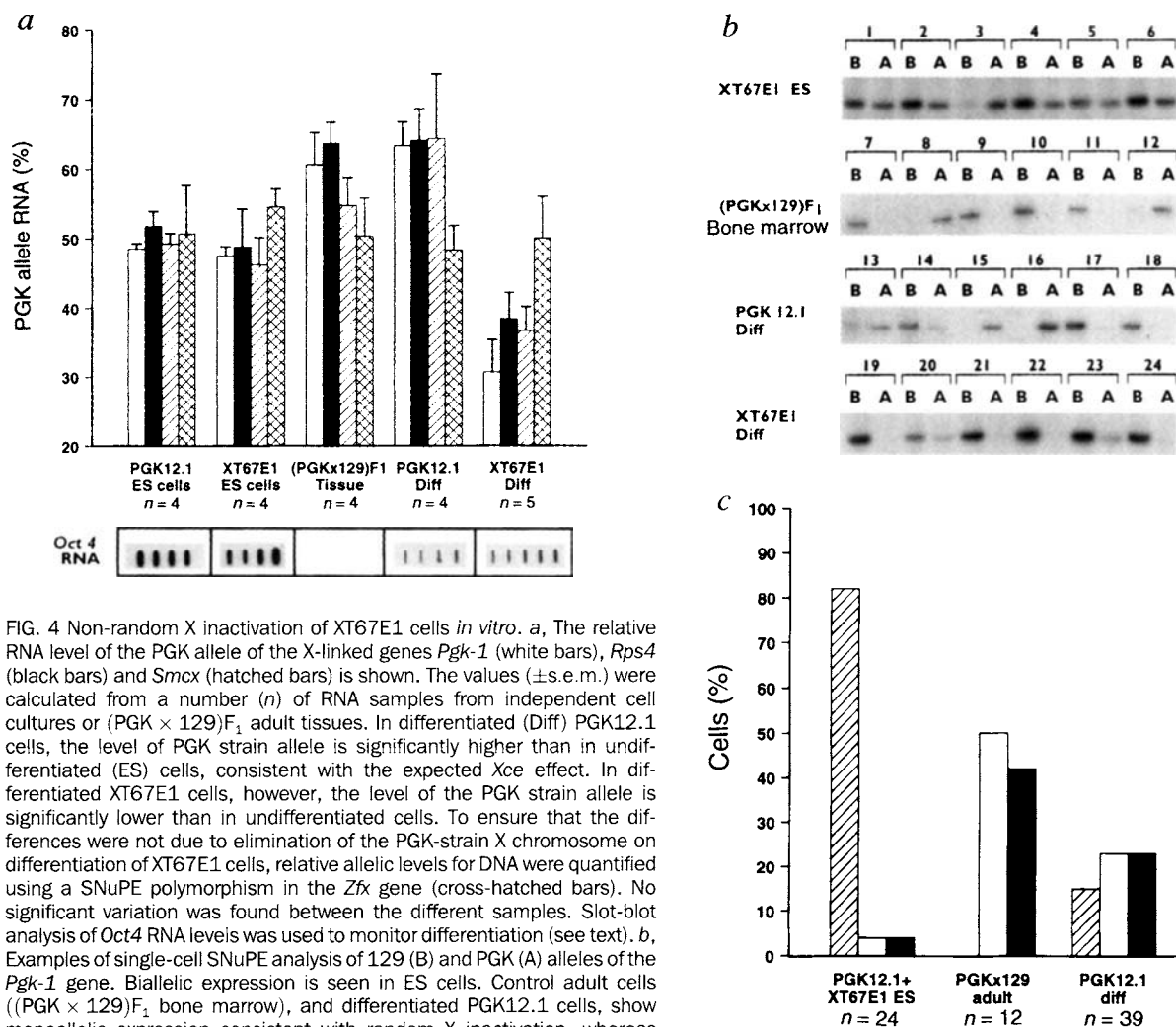


FIG. 4 Non-random X inactivation of XT67E1 cells *in vitro*. **a**, The relative RNA level of the PGK allele of the X-linked genes *Pgk-1* (white bars), *Rps4* (black bars) and *Smcx* (hatched bars) is shown. The values ($\pm$ s.e.m.) were calculated from a number (*n*) of RNA samples from independent cell cultures or (PGK x 129)F₁ adult tissues. In differentiated (Diff) PGK12.1 cells, the level of PGK strain allele is significantly higher than in undifferentiated (ES) cells, consistent with the expected Xce effect. In differentiated XT67E1 cells, however, the level of the PGK strain allele is significantly lower than in undifferentiated cells. To ensure that the differences were not due to elimination of the PGK-strain X chromosome on differentiation of XT67E1 cells, relative allelic levels for DNA were quantified using a SNUPE polymorphism in the *Zfx* gene (cross-hatched bars). No significant variation was found between the different samples. Slot-blot analysis of *Oct4* RNA levels was used to monitor differentiation (see text). **b**, Examples of single-cell SNUPE analysis of 129 (B) and PGK (A) alleles of the *Pgk-1* gene. Biallelic expression is seen in ES cells. Control adult cells ((PGK x 129)F₁ bone marrow), and differentiated PGK12.1 cells, show monoallelic expression consistent with random X inactivation, whereas differentiated XT67E1 cells show monoallelic expression of the 129 allele only. **c**, Summary of quantified single-cell analysis. Cells showing 50 $\pm$ 10% levels of both alleles were scored as having no inactive X chromosome (hatched bars). Cells with > 80% 129 (white bars) or PGK (black bars) allele were scored as having undergone X inactivation (see Methods).

METHODS. Cell culture, *Oct4* analysis and SNUPE analysis were as described in Fig. 1. New polymorphisms were found for the *Rps4* gene⁴¹ (position 399; PGK strain, G; 129 strain, A), the *Smcx* gene³² (position 228 of the published partial cDNA sequence; PGK strain, C; 129 strain, G) and the *Zfx* gene⁴² (position 4311; PGK strain, T; 129 strain, A). Primers for *Rps4*; Rps4-323; AGGGTCGCTTTGCTGTCA and Rps4-885; AGTTTCAC-CATGCTGTTA (*Rps4* RT-PCR, 30 cycles of 94 °C 1', 52 °C 1', 72 °C 1') and Rps4-358; CAGGTGCGGGATTCCTTTGTGCC (*Rps4* SNUPE, 1 cycle of 94 °C 1', 62 °C 2', 72 °C 1'). For *Smcx*; Smcx-55; TGGTTCCTTGCTACGCTCT-CACTA and Smcx-889; CAGCCGCCAAAGCTCCTTCTAC (*Smcx* RT-PCR, 30 cycles of 94 °C 1', 55 °C 1', 70 °C 1'), and Smcx-204; CTGTGACGCTTCAAGTCAACA (*Smcx* SNUPE, 1 cycle of 94 °C 1', 59 °C 2', 72 °C 1'), and for *Zfx*; Zfx-4068; GTGAACGGGTGCTAAGGAGGACT, and Zfx-4458; TAACGGCACGAAGGATGG (*Zfx* PCR, 30 cycles of 94 °C 1', 55 °C 1', 72 °C 1'), and Zfx-4312; AAAGACATCTTAACAAAATCTTAATA (*Zfx* SNUPE, 1 cycle of 94 °C 1', 45 °C 2', 72 °C 1'). Single-cell analysis was

performed by diluting trypsinized single-cell suspensions at 10⁴ cells per ml and pipetting 0.2 μ l microdrops into 96-well plates. Lysis solution (10 μ l; 0.05% NP40, 1 mg ml⁻¹ glycogen, 10 μ l ml⁻¹ RNasin) was added to wells containing a single cell. Negative PCR controls were provided by microdrops without cells. Samples were ethanol precipitated and redissolved in 5 μ l of RT mix. After RT (using gene-specific primer, *Pgk750* (ref. 24)), PCR mix (65 μ l) was added directly to samples, and PCR (40 cycles) was performed with primers *Pgk750* and *Pgk402* (CCTCCGCTTCATGTAGAGGAAGA). SNUPE was performed as described previously²⁴. Control bone-marrow cells were cultured as described previously⁴³. Several differentiated PGK12.1 and XT67E1 single cells showed near-complete monoallelic expression, rather than complete monoallelic expression (see, for example, Fig. 4b, cells 13, 14, 17 and 23), presumably because they are contaminated by trace amounts of cell debris and RNA owing to high levels of apoptosis in differentiated cultures, and/or cell damage arising from the preparation of single-cell suspensions. Also, differentiated cultures will contain cells that have recently undergone X inactivation, and therefore have trace amounts of RNA from the inactive allele still present. We therefore scored cells with > 80% monoallelic expression as having undergone X inactivation.

with random X inactivation subject to the predicted *Xce* effect. Gene targeting by homologous recombination was used to create a 7-kb deletion of the 129 derived *Xist* gene resulting in a null allele. On differentiation *in vitro*, targeted cells undergo X inactivation, indicating that the counting function of the *Xic* is unaffected. However, unlike parent PGK12.1 cells, targeted cells differentiated *in vitro* or *in vivo* display complete non-random X inactivation of the PGK strain X chromosome bearing

FIG. 5 XT67E1 cells show complete non-random X inactivation *in vivo*. Chimaeric embryos were made by aggregation of eight-cell CD1 host embryos with either XT67E1 or PGK12.1 donor ES cells. RNA was prepared from embryos 10.5–12.5 d.p.c.. SNUPE analysis was performed using polymorphisms in the *Xist*, *Pgk-1* and *Smcx* genes. The presence of a Y chromosome indicates the chromosomal sex of the host embryo. This was determined by PCR for the *Zfy* gene from yolk sac DNA as described previously²⁸. In an XY host embryo, the PGK12.1 cells expressed both the PGK and 129 strain *Xist* alleles, and the PGK and 129/CD1 strain *Pgk-1* alleles, consistent with random X inactivation. Targeted XT67E1 cells, however, expressed the PGK *Xist* allele and only the 129/CD1 *Pgk-1* allele in XX or XY host embryos, indicating that all donor cells have an inactive PGK-derived X chromosome. Both alleles of the *Smcx* gene were expressed in all chimaeric embryos, consistent with this gene escaping X inactivation.

METHODS. Aggregation chimaeras were prepared with eight-cell CD1 host embryos, essentially as described previously⁴⁴, except that ES cell clumps were produced by a short trypsinization, stopped by addition of M16 with 10% fetal calf serum, and aggregations were performed by laying small (4–12 cell) ES cell clumps onto compacting eight-cell embryos placed in microwells. After overnight incubation in M16 (5% CO₂, 37 °C, 100% humidity), aggregates were transferred in M16 with 10% fetal calf serum to the uterine horns of day 2.5 pseudopregnant CD1 foster mothers. Chimaeric embryos were collected from recipient mothers at 10.5–

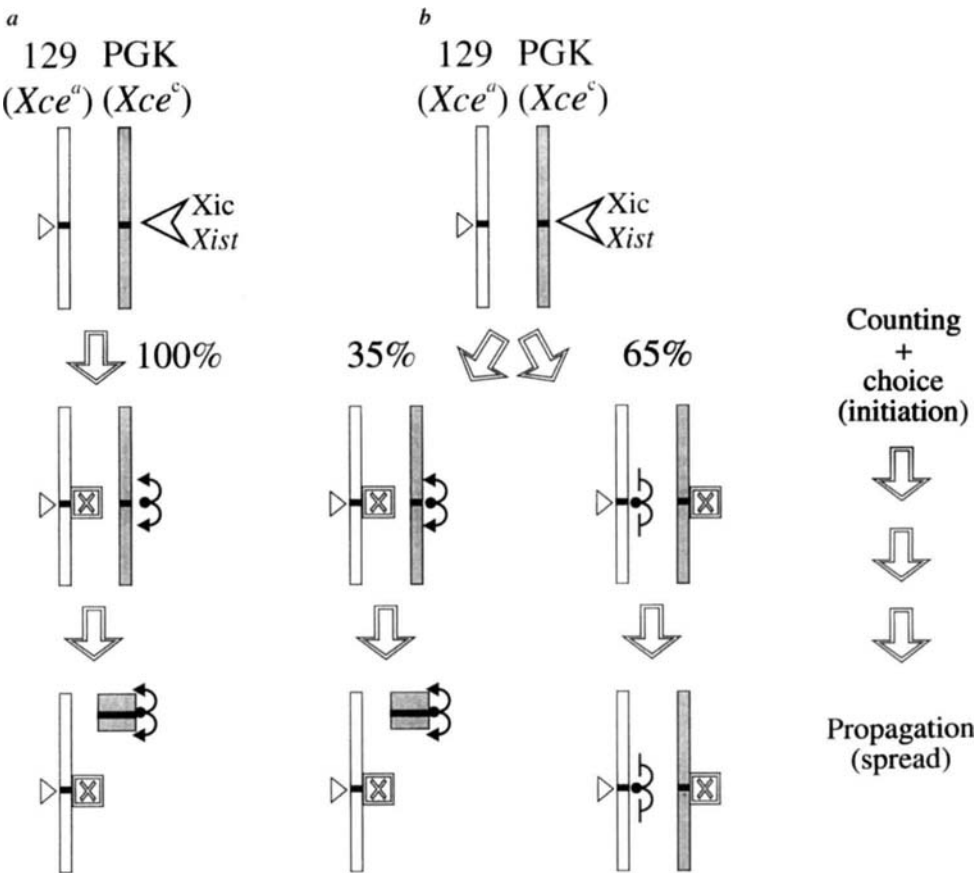
the intact *Xist* allele. These results show that *Xist* is required *in cis* for X inactivation to occur.

Non-random X inactivation in XT67E1 cells may be occurring as a result of either primary or secondary mechanisms, as illustrated in the models in Fig. 6. Primary non-random X inactivation would result if the *Xist* deletion disrupts the choice function of the *Xic*, that is, the XT67E1 cells always inactivate the PGK X chromosome bearing the non-targeted *Xist* locus (Fig. 6a). The

Allele	PGK 129/CD1	PGK 129/CD1	PGK 129/CD1	PGK 129/CD1
<i>Xist</i>				
<i>Pgk-1</i>				
<i>Smcx</i>				
Host embryo	XY	XX	XX	XY
Cell line	XT67E1			PGK12.1

12.5 d.p.c. SNUPE analysis of RNA from chimaeric embryos was as described in Figs 1 and 4.

FIG. 6 Primary and secondary non-random X inactivation models for targeted XT67E1 cells. These models are based on the proposal that each cell has a limited amount of a blocking factor (indicated by open squares with a cross), sufficient to bind and repress a single *Xic*⁴. *In cis* propagation of X inactivation is triggered from unblocked *Xics* on differentiation of cells from the totipotent embryonic lineage. This proposal is consistent with the observation that one X chromosome remains active per diploid autosome set in sex-chromosome aneuploid⁴⁵ and tetraploid cells⁴⁶. a, Primary non-random X inactivation results if the deleted *Xist* allele on the 129 X chromosome (white bar) is blocked in all cells. *Xist* expression (curved arrows) proceeds on the PGK X chromosome (grey bar), resulting in an inactive PGK X chromosome (grey square) in all cells. b, Secondary non-random X inactivation results if either *Xist* allele can be blocked, but cells in which the normal allele is blocked fail to express *Xist* and do not propagate X inactivation *in cis* from the deleted allele (blunt curves). Cells which block the deleted *Xist* allele, however, express *Xist* from the PGK X chromosome and go on to X inactivate *in cis*. The proportion of these two cell populations is governed by the *Xce* effect as indicated, but differentiated cells with two active X chromosomes will be selected against as a result of failure to dosage compensate the X chromosome.



Xist deletion would then be equivalent to a very strong *Xce* allele. Secondary non-random X inactivation would result if the counting and choice functions of the *Xic* are unaffected by the *Xist* deletion, and cells that elect to inactivate the 129 X chromosome (bearing the targeted *Xist* locus) fail to X inactivate *in cis*. Cells that elect to inactivate the PGK X chromosome (bearing an intact *Xist* locus), however, will undergo X inactivation (Fig. 6b).

Although we cannot state definitively which mechanism is operating in XT67E1 cells, our data favour secondary non-random X inactivation. Analysis of *Oct4* expression indicates *in vitro* differentiation levels greater than 80%, both for PGK12.1 and XT67E1 cells (Fig. 4a). On this basis, primary non-random X inactivation (Fig. 6a) would lead to predominant expression of the 129 allele of X-linked genes in differentiated cells. We observed 129 allele values of 65%, 63% and 67% for the *Pgk-1*, *Rps4* and *Smcx* genes, respectively (Fig. 4a). These values are more consistent with the secondary non-random X inactivation model (Fig. 6b); according to this model the proportion of cells inactivating the PGK- and 129-derived X chromosomes will be governed by the *Xce* effect, as in control cells. Thus, in differentiated XT67E1 cells, approximately 65% of cells would inactivate from the targeted 129 *Xist* allele, and therefore go on to differentiate into cells with both X chromosomes active. The predicted level of 129 strain RNA for an X-linked gene undergoing X inactivation would therefore be 60.6%, compared to the 35% level predicted and observed in control PGK12.1 cells. That we see slightly higher levels of the 129 allele than predicted by this model could reflect selection *in vitro* against differentiated cells with two active X chromosomes. The single-cell analysis (Fig. 4c) and the late-replication analysis (Fig. 3b) indicate that the proportion of XX active cells is higher in differentiated XT67E1 cultures than in differentiated PGK12.1 cultures, consistent with the secondary non-random X inactivation model. In both cases this difference is

less than predicted in the model, but again this could be accounted for by selection *in vitro* against differentiated cells with two active X chromosomes. The complete non-random X inactivation observed *in vivo* is consistent either with primary or secondary non-random X inactivation, as T(X;16)16H/+ cells which are functionally disomic for the proximal half of the X chromosome are rapidly selected against after the onset of X inactivation^{34,35}.

It is possible that the non-random X inactivation observed results from an unrelated, spontaneous cell-lethal mutation on the 129 X chromosome, but such a mutation would need to result in extremely rapid cell death *in vitro*, as no single cells with an inactive PGK X chromosome were seen (Fig. 4c). We therefore consider this possibility highly unlikely.

In summary, our data demonstrate that an intact *Xist* gene is required *in cis* for X inactivation to occur. We propose that the phenotype of XT67E1 cells results from secondary non-random X inactivation of the chromosome bearing the non-deleted *Xist* allele, and that the counting and choice functions of the *Xic* are not affected by the deletion. If correct, this would support our previous suggestion that the primary decision determining how many and which X chromosomes to inactivate involves the choice of which, if any, *Xist* alleles are expressed, and that subsequent propagation of X inactivation *in cis* results from expression of the *Xist* gene^{16,18}. We predict that the sequences required for counting and choice reside in regulatory elements of the *Xist* gene. These elements may lie in the region 5' of the gene, but could also be located 3' to *Xist*, consistent with genetic mapping studies which suggest that *Xce* lies distal to *Xist*³⁶. Further studies will be required to determine the mechanism of *Xist* function, as our data are consistent both with models that propose *Xist* RNA as the functional *cis*-inactivating signal^{16,17} and those that propose the underlying chromatin structure of the active *Xist* locus as the primary determinant of *cis* inactivation¹⁶. □

Received 18 August; accepted 24 November 1995.

1. Lyon, M.F. *Nature* **190**, 372–373 (1961).
2. Rastan, S. *Curr. Opin. Genet. Dev.* **4**, 292–297 (1994).
3. Rastan, S. & Robertson, E. J. *J. Embryol. exp. Morph.* **90**, 379–388 (1985).
4. Rastan, S. *J. Embryol. exp. Morph.* **78**, 1–22 (1983).
5. Cattanach, B. M. & Isaacson, J. H. *Genetics* **57**, 331–346 (1967).
6. Cattanach, B. M., Pollard, C. E. & Perez, J. N. *Genet. Res.* **14**, 223–235 (1969).
7. Cattanach, B. M., Perez, J. N. & Pollard, C. E. *Genet. Res.* **15**, 183–195 (1970).
8. Rastan, S. *Genet. Res.* **40**, 139–147 (1982).
9. Russell, L. B. & Montgomery, C. S. *Genetics* **64**, 281–312 (1970).
10. Cattanach, B. M. *Genet. Res.* **23**, 291–306 (1974).
11. Brown, C. J. & Willard, H. F. *Nature* **368**, 154–156 (1994).
12. Brown, C. J. *et al. Nature* **349**, 82–84 (1991).
13. Brown, C. J. *et al. Nature* **349**, 38–44 (1991).
14. Brockdorff, N. *et al. Nature* **351**, 329–331 (1991).
15. Borsani, G. *et al. Nature* **351**, 325–329 (1991).
16. Brockdorff, N. *et al. Cell* **71**, 515–526 (1992).
17. Brown, C. J. *et al. Cell* **71**, 527–542 (1992).
18. Kay, G. F. *et al. Cell* **72**, 171–182 (1993).
19. Kay, G. F., Barton, S. C., Surani, M. A. & Rastan, S. *Cell* **77**, 639–650 (1994).
20. McCarrey, J. R. & Dilworth, D. D. *Nature Genet.* **2**, 200–203 (1992).
21. Salido, E. C., Yen, P. H., Mohandas, T. K. & Shapiro, L. J. *Nature Genet.* **2**, 196–199 (1992).
22. Green, M. C. in *Genetic Variants and Strains of the Laboratory Mouse* 2nd edn (eds Lyon, M. F. & Searle, A. G.) 282 (Oxford Univ. Press, 1989).
23. Norris, D. P. *et al. Cell* **77**, 41–51 (1994).
24. Singer Sam, J., LeBon, J. M., Dai, A. & Riggs, A. D. *PCR Meth. Appl.* **1**, 160–163 (1992).
25. Singer Sam, J., Chapman, V., LeBon, J. M. & Riggs, A. D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10469–10473 (1992).
26. Buzin, C. H., Mann, J. R. & Singer Sam, J. *Development* **120**, 3529–3536 (1994).

27. Pamieri, S. L., Peter, W., Hess, H. & Scholer, H. R. *Dev. Biol.* **166**, 259–267 (1994).
28. Mansour, L. S., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348–353 (1988).
29. Dutilleul, B. & Fosse, A.-M. *Ann. Genet.* **17**, 207–211 (1974).
30. Tada, T., Tada, M. & Takagi, N. *Development* **119**, 813–821 (1993).
31. Takagi, N. & Martin, G. R. *Dev. Biol.* **103**, 425–433 (1984).
32. Agulnik, A. I. *et al. Hum. molec. Genet.* **3**, 879–884 (1994).
33. Wu, J. *et al. Nature Genet.* **7**, 491–496 (1994).
34. McMahon, A. & Monk, M. *Genet. Res.* **41**, 69–83 (1983).
35. Takagi, N. & Abe, K. *Development* **109**, 189–201 (1990).
36. Simmler, M. C., Cattanach, B. M., Rasberry, C., Rougeulle, C. & Avner, P. *Mamm. Genome* **4**, 523–530 (1993).
37. Robertson, E. J. in *Teratocarcinomas and Embryonic Stem Cells, a Practical Approach* (ed Robertson, E. J.) 71–112 (IRL, Oxford, 1987).
38. Scholer, H. R., Rupperts, S., Suzuki, N., Chowdury, K. & Gruss, P. *Nature* **344**, 435–439 (1990).
39. Brockdorff, N., Kay, G., Cattanach, B. M. & Rastan, S. *Mamm. Genome* **1**, 152–157 (1991).
40. Ramirez-Solis, S. *et al. Anal. Biochem.* **201**, 331–335 (1992).
41. Zinn, A. R. *et al. Genomics* **11**, 1097–1101 (1991).
42. Mardon, G. *et al. Molec. cell. Biol.* **10**, 681–688 (1990).
43. Ball, T. C., Hirayama, F. & Ogawa, M. *Blood* **85**, 3086–3092 (1995).
44. Nagy, A. & Rossant, J. in *Gene Targeting, a Practical Approach* (ed Joyner, A. L.) 147–178 (IRL, Oxford, 1993).
45. Grumbach, M. M., Morishima, A. & Taylor, J. H. *Proc. natn. Acad. Sci. U.S.A.* **49**, 581–589 (1963).
46. Webb, S., de Vries, T. J. & Kaufman, M. H. *Genet. Res.* **59**, 205–214 (1992).

ACKNOWLEDGEMENTS. We thank N. Takagi for advice on late-replication assays; A. Fisher for help with bone-marrow tissue culture; and R. Beddington, V. Episkopou and members of the Section of Comparative Biology for valuable discussions. This work was supported by the Medical Research Council of Great Britain.

The redshift of the gravitational lens of PKS1830–211 determined from molecular absorption lines

T. Wiklind* & F. Combes†

* Onsala Space Observatory, S-43992 Onsala, Sweden

† DEMIRM, Observatoire de Paris, 61 Avenue de l'Observatoire, F-75014 Paris, France

GRAVITATIONAL lensing can be used to derive fundamental cosmological parameters, provided that the redshift of the lens and the structure of its potential well are known¹. The radio source PKS1830–211—a lensed quasar^{2–5} whose components are separated by about 1 arcsecond—is ideal for such an examination, but until recently the redshift of the lens has been unknown. Here we report the detection and identification of 12 molecular absorption lines in the spectrum of PKS1830–211; the lines originate in the lensing source, which is probably a spiral galaxy, at a redshift of ~ 0.89 . Only one of the lensed components is covered by the intervening molecular gas; when combined with the strong variability in the quasar, this creates a unique opportunity to measure the difference in the light travel time for the two components, which depends in part on the Hubble constant and the cosmic deceleration parameter.

The two major radio-components in the gravitational lens PKS1830–211 have almost identical substructure^{2,3}. Within the framework of the gravitational lensing hypothesis, the two components are viewed as two parity-reversed images of a background quasar consisting of a core, a linearly polarized knot and an extended jet-like feature^{2–5}. A third demagnified image has been tentatively identified^{2,4}. Although PKS1830–211 is located close to the Galactic Centre (galactic longitude $l = 12.2^\circ$, latitude $b = -5.7^\circ$), the galactic extinction is moderate, with $A_V \approx 2.7$ mag (ref. 2). Nevertheless, deep imaging in the B, V, R, I and K bands^{2,3,6} has failed to detect an optical counterpart, with an upper limit to the B-band magnitude of ~ 23 . A search for redshifted 21-cm H I absorption was also unsuccessful⁷, owing to man-made interferences at these low frequencies.

Previous observations of three galaxies at redshifts $z = 0.25$, 0.67 and 0.68 have shown that absorption of molecular rotational lines in the millimetre band is a very sensitive method of studying the physical and chemical conditions of molecular gas at large distances^{8–10}. These observations have been done for objects where the redshift of the intervening galaxy was known beforehand from optical spectroscopy, even if the background source itself is invisible due to heavy obscuration. However, the high sensitivity of molecular absorption, even to relatively modest amounts of gas in the line of sight, makes it possible to search for lines of unknown redshift in the millimetre-wave band.

In June 1995 we used the 15-m SEST telescope at La Silla to search for molecular absorption lines towards PKS1830–211. As the separation of the compact components and the angular extent of the ring is ~ 1 arcsec, the line of sight towards the background source should pass within a few kiloparsecs of the centre of the hypothesized lens, making absorption from an intervening molecular cloud possible. For redshifts $z \leq 1.2$ we would have one or more of the CO($1 \leftarrow 0$), CO($2 \leftarrow 1$), HCN($2 \leftarrow 1$), HCO⁺($2 \leftarrow 1$) and HNC($2 \leftarrow 1$) lines in the 3- and 2-mm band. From our experience with the three previously observed molecular absorption line systems^{8–10}, we know that these transitions are likely to show saturated absorption even in the case of diffuse molecular gas. The continuum flux at 3-mm wavelength was 2.1 Jy at the time of the observation, and is known to be variable. The

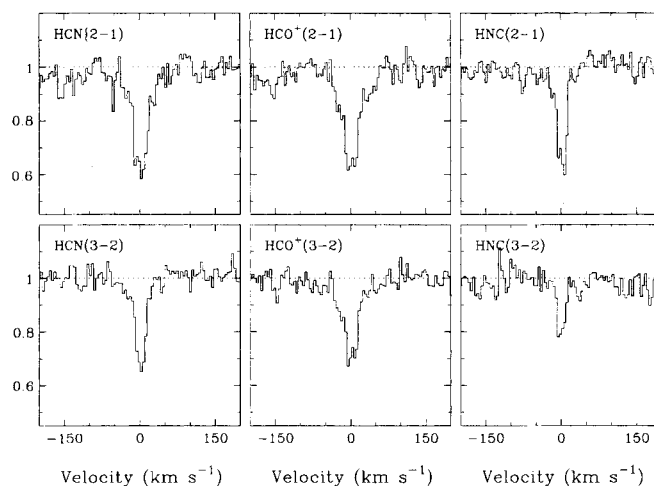


FIG. 1 The $J = 2 \leftarrow 1$ and $J = 3 \leftarrow 2$ absorption spectra of HCN, HCO⁺ and HNC. The continuum level has been normalized to unity for all transitions. These lines are likely to be saturated despite not reaching the zero continuum level.

SEST telescope is equipped with low-noise SIS receivers with an instantaneous bandwidth of 1 GHz and has a remote tuning facility, making rapid changes of the observed frequency possible. This allowed us to search over 14 GHz of the 3-mm band within a few hours. When the first molecular absorption line was detected we searched at selected frequencies in the 3- and 2-mm bands to establish the identity of the line, and obtained the redshift of the lensing galaxy as $z = 0.88582 \pm 0.00001$. In total, we observed 12 transitions: the $J = 2 \leftarrow 1$ and $J = 3 \leftarrow 2$ lines of HCN, HCO⁺, HNC, N₂H⁺ and H¹³CO⁺, and the $J = 3 \leftarrow 2$ and $J = 4 \leftarrow 3$ lines of CS. The CO lines are not observable in the 3- and 2-mm band at this redshift. The observed absorption profiles are shown in Figs 1 and 2.

The two images of the core of PKS1830–211 are situated on opposite sides of the centre of the lensing galaxy⁴, with the northeastern (NE) image at 3.8 kpc and the southwestern (SW) image at 1.8 kpc from the lens centre (assuming that the Hubble constant $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and the deceleration parameter $q_0 = 0.5$). If both core images are covered by molecular gas, we expect two absorption profiles with a velocity separation of a few hundred km s^{-1} . However, the absorption consists of a single component with a maximum width of $\sim 30 \text{ km s}^{-1}$. Moreover, the absorption profiles do not reach the zero level, as expected for an optically thick molecular cloud covering the entire background source. As the H¹³CO⁺($2 \leftarrow 1$) line is detected and because all the $J = 2 \leftarrow 1$ lines of HCN, HCO⁺ and HNC reach a similar depth, the absorption is nevertheless likely to be saturated. The mere detection of the H¹³CO⁺($2 \leftarrow 1$) line means that the optical depth of the HCO⁺($2 \leftarrow 1$) line is $\gg 1$. This implies that only part of the background continuum is covered by molecular gas, but that this gas must be optically thick. The fraction f_c of the continuum source area that is obscured by optically thick molecular gas can be expressed as $f_c = T_{\text{abs}}/T_c$, where T_{abs} is the depth of the absorption line measured from the continuum level T_c (refs 8, 9). We find that 36% of PKS1830–211 is obscured by molecular gas.

The flux ratio between the two lensed images of the compact core represents different point-image magnifications. In PKS1830–211 the flux ratio appears to increase with frequency. Because each radio-component consists of an unresolved image of a flat-spectrum core, where the flux is more or less independent of frequency, and a steep-spectrum knot, where the flux decreases with increasing frequency, the varying flux ratio can be ascribed to the core becoming more dominant at higher frequencies⁴. In fact,

at millimetre wavelengths the flux from the extended structure has decreased to negligible levels and the observed continuum is completely dominated by the two compact cores. Separating the intrinsic core and knot magnification ratios in the very-long-baseline-interferometer images obtained at low frequencies, using the values given in ref. 3, we derive a magnification ratio for the NE and SW core of 1.78 ± 0.10 . If the molecular gas only covers one of the compact cores, the point-image magnification ratio R is directly given by $R = (T_c/T_{\text{abs}}) - 1 = f_c^{-1} - 1$. Our molecular absorption line data implies a point-image magnification ratio of 1.75 ± 0.15 , where the uncertainties on the total continuum flux and on the depth of the absorption profiles have been taken into account. From this we conclude that the SW core is probably obscured by molecular gas, while the NE component is not.

PKS1830–211 is known¹¹ to have changed its millimetre continuum flux by a factor of two on a timescale of ~ 8 months. As scintillation is negligible at millimetre wavelengths, this variation is likely to be intrinsic to the background source. The saturated $J = 2 \leftarrow 1$ lines of HCN, HCO^+ and HNC can therefore be used as a sensitive probe of the time-delay between the SW and NE images of the core. An accurate measurement of the time-delay and a good model of the potential of the lensing galaxy can give an estimate of the distance of the lens, involving both q_0 and H_0 (ref. 1). The time-delay can also be used as an aid in refining the lens potential. Owing to the fortunate configuration of the molecular gas in the lensing galaxy, the point-image magnification ratio R is directly obtained by a relatively short integration on a single-dish telescope. The saturated molecular absorption lines are therefore our best way to easily and accurately monitor the point-image magnification ratio R of PKS1830–211 and, hence, to obtain the time-delay. The expected time-delay between the SW and NE cores depends on the mass and the shape of the potential of the lensing galaxy. If the mass of the lensing galaxy is $(2\text{--}6) \times 10^{11} M_\odot$, then the lens model of ref. 4 and an image separation of 1.0 arcsec gives a time-delay of 3–7 weeks, with the SW component leading.

Molecular gas is usually found in spiral galaxies, although a significant number of nearby lenticular and elliptical galaxies are known to contain a molecular interstellar-medium component as well¹². Nevertheless, a galaxy at $z = 0.89$ containing molecular gas with a substantial amount of heavy elements at a galactocentric distance of ~ 2 kpc, is most likely a spiral. The redshift of $z = 0.89$ corresponds to a Universe which is 38% of its present age, which translates to a lookback time of 5.4 Gyr ($H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.5$). It is therefore not surprising that we find an interstellar medium containing molecules made up of processed elements such as carbon, nitrogen and oxygen. It is, however, interesting to note that the conditions for stellar formation do not

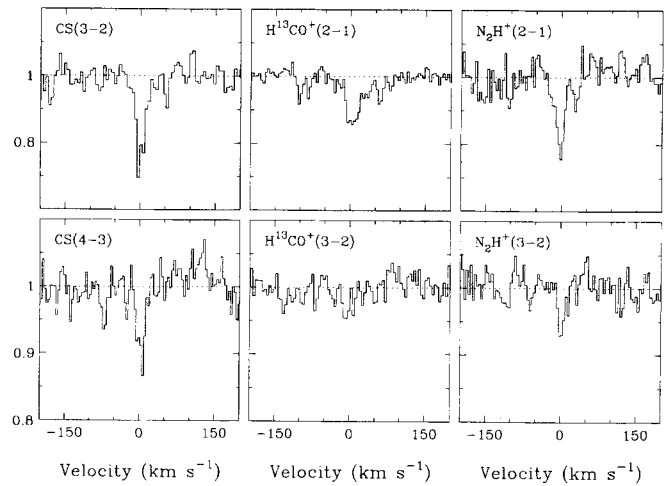


FIG. 2 The absorption spectra of CS, H^{13}CO^+ and N_2H^+ . The continuum level has been normalized to unity for all transitions. These lines are unsaturated. For clarity, the spectra have different scales from those of Fig. 1.

appear to be different at this epoch from what we find in our own Galaxy. This enables us to put faith in models of galactic evolution based on local star formation processes. Unfortunately it will be difficult, or impossible, to obtain optical data on the lensing galaxy. A typical L^* galaxy ($L_B = 10^{10} L_\odot$) at $z = 0.89$ would have an apparent blue magnitude of ~ 23.6 , which is at the limit of optical searches towards PKS1830–211 (ref. 6). Accounting for the redshifted spectral energy distribution with a K-correction typical for spiral galaxies, the apparent magnitude will increase by 1–2 mag. Together with an estimated Galactic extinction of $A_B = 3.5$ mag, an optical counterpart to the lensing galaxy will have an apparent B-band magnitude of $\sim 28\text{--}29$. Very deep images in the near-infrared, where the extinction is much lower, may be a possible way to obtain data on the stellar population of the lensing galaxy.

The saturated lines of HCN, HCO^+ and HNC, give upper limits to the rotational temperature T_{rot} and lower limits to the column densities. The unsaturated profiles directly give T_{rot} and corresponding column densities, but because these lines are weaker, the associated uncertainties are still relatively large. For HCN, HCO^+ and HNC we estimate T_{rot} to be less than 7.6, 8.4 and 6.0 K, respectively. For HNC, which appears to have the least saturated profile, T_{rot} is only slightly higher than the expected temperature of the cosmic microwave background radiation T_{CMB} of 5.16 K. The unsaturated CS, H^{13}CO^+ and N_2H^+ lines give $T_{\text{rot}} \approx 4.0$ K. Given the 3σ errors associated with these values (see Table 1), they are

TABLE 1 Molecular-gas properties

	ν_{obs} (GHz)		$\int \tau_\nu dV$ (km s^{-1})			T_{rot} (K)	$\log N^\dagger$ (cm^{-2})
			$J = 2 \leftarrow 1$	$J = 3 \leftarrow 2$	$J = 4 \leftarrow 3$		
Observed							
CS	77.934	103.910		27.0 ± 4.8	6.9 ± 3.3	$4.0^{+1.8}_{-1.4}$	$14.682^{+0.071}_{-0.085}$
H^{13}CO^+	92.006	138.007	17.5 ± 3.3	3.6 ± 2.4		$4.0^{+1.9}_{-1.4}$	$13.746^{+0.084}_{-0.081}$
HCN	93.997	140.993	$\geq 82.3 \pm 6.6$	48.8 ± 5.4		$\leq 7.6^{+1.3}_{-1.0}$	$\geq 14.499^{+0.034}_{-0.036}$
HCO^+	94.588	141.879	$\geq 97.1 \pm 6.3$	59.2 ± 6.0		$\leq 8.4^{+1.4}_{-1.1}$	$\geq 14.466^{+0.029}_{-0.031}$
HNC	96.152	144.225	$\geq 53.3 \pm 5.7$	21.5 ± 5.4		$\leq 6.0^{+1.5}_{-1.2}$	$\geq 14.405^{+0.044}_{-0.049}$
N_2H^+	98.814	148.218	18.7 ± 4.2	3.4 ± 3.0		$4.1^{+2.4}_{-2.1}$	$13.730^{+0.088}_{-0.111}$
Estimated							
$\text{CO}^\ddagger$							18.5
$\text{H}_2^\ddagger$							22.5

$\int \tau_\nu dV$ is the optical depth integrated over velocity; N is the column density.

* Observed frequency: $\nu_{\text{obs}} = \nu_0/(1+z)$, where ν_0 is the rest frequency of the transition.

† Estimated assuming a Galactic abundance ratio $N(\text{CO}) \approx 10^4 N(\text{HCN}) \text{ cm}^{-2}$.

‡ Estimated assuming a Galactic abundance ratio $N(\text{H}_2) \approx 10^4 N(\text{CO}) \text{ cm}^{-2}$.

still compatible with $T_{\text{CMB}} = 5.16$ K. The actual kinetic temperature can be much higher. The low value of T_{rot} is an indication that the H_2 density is not high enough to rotationally excite the molecules. In fact, the very low values of T_{rot} obtained for the non-saturated molecules show that the excitation of the molecular lines is dominated by the cosmic background radiation. From our molecular absorption line data we derive a conservative upper limit to the temperature of the cosmic background radiation at $z = 0.88582$ of $T_{\text{CMB}} \leq 6.0$ K.

Using $T_{\text{rot}} = 6.0$ K, we estimate the molecular column densities presented in Table 1. Although three of these represent lower limits, they are nevertheless rather high and suggest that the absorbing molecular gas resides in a large cloud complex. We estimate the CO and H_2 column densities in the absorbing gas at $z = 0.88582$ to be at least 3×10^{18} and $3 \times 10^{22} \text{ cm}^{-2}$, respectively. These values are derived using Galactic abundance ratios between HCN, CO and H_2 . Should the metallicity in the galaxy at $z = 0.89$ be lower than in the Milky Way, the column density of H_2 will be even higher. Because the excitation of the molecules is dominated by the cosmic background radiation, the volume density is mod-

erate despite the high H_2 column density. This suggests that the absorbing gas is made up of several diffuse gas clouds, possibly corresponding to the intercloud medium of a giant molecular complex. The low H_2 density accounts for the low collisional excitation, while the width of the absorption lines as well as the high column densities implies the presence of a massive complex along the line of sight. $\square$

Received 23 August; accepted 15 November 1995.

1. Narayan, R. *Astrophys. J.* **378**, L5–L9 (1991).
2. Subrahmanyan, R., Narasimha, D., Rao, A. P. & Swarup, G. *Mon. Not. R. astr. Soc.* **246**, 263–272 (1990).
3. Jauncey, D. L. et al. *Nature* **352**, 132–134 (1991).
4. Nair, S., Narashima, D. & Rao, A. P. *Astrophys. J.* **407**, 46–59 (1993).
5. Kochanek, C. S. & Narayan, R. *Astrophys. J.* **401**, 461–473 (1992).
6. Djorgovski, S. et al. *Mon. Not. R. astr. Soc.* **257**, 240–244 (1992).
7. McMahon, P. M., Moore, C., Hewitt, J. N., Rupen, M. P. & Carilli, C. *Bull. Am. astr. Soc.* **25**, 1307 (1993).
8. Wiklind, T. & Combes, F. *Astr. Astrophys.* **286**, L9–L12 (1994).
9. Wiklind, T. & Combes, F. *Astr. Astrophys.* **299**, 382–388 (1995).
10. Wiklind, T. & Combes, F. *Astr. Astrophys.* (submitted).
11. Steppe, H. et al. *Astr. Astrophys. Suppl. Ser.* **102**, 611–635 (1993).
12. Wiklind, T., Combes, F. & Henkel, C. *Astr. Astrophys.* **297**, 643–659 (1995).

Experimental simulations of the photodecomposition of carbonates and sulphates on Mars

L. M. Mukhin^{*†}, A. P. Koscheev[‡], Yu. P. Dikov[§], J. Huth^{||} & H. Wänke^{||}

^{*} Embassy of the Russian Federation, 2650 Wisconsin Avenue, NW, Washington DC 20007, USA

[†] Space Research Institute, Moscow, Russia

[‡] Active Particle Adsorption Laboratory, Karpov Institute of Physical Chemistry, 10, Ul. Obukha 10, 103064 Moscow E-64, Russia

[§] Institute of Ore Deposits, Petrography, Mineralogy and Geochemistry, Russian Academy of Sciences, Staromonetny per. 35, 109017 Moscow, Russia

^{||} Max-Planck-Institut für Chemie, Abteilung Kosmochemie, Saarstrasse 23, D-55122 Mainz, Germany

THERE is indirect spectroscopic evidence for the presence of sulphates and carbonates on the martian surface^{1–5}, and such minerals are also found in SNC meteorites⁶, which are thought to be of martian origin. But although carbonates are expected to be abundant in the martian regolith^{7–9}, attempts to detect them directly have been unsuccessful^{10,11}. Here we report laboratory studies of the decomposition of calcium carbonate and magnesium sulphate under ultraviolet irradiation, which mimic the conditions under which photodecomposition of surface minerals by solar ultraviolet light might occur on Mars. We find that, even for a low abundance⁶ of carbonate minerals in the martian regolith, the rate of CO_2 release due to photodecomposition is higher than the rate of CO_2 loss from the atmosphere by solar-wind-induced sputtering processes^{12–15}, making this process a potential net source of atmospheric CO_2 over time. SO_2 is also released from the sulphate, albeit more slowly. The rate of carbonate degradation is high enough to explain the apparent absence of these compounds at the martian surface.

We investigated the stability of calcium carbonate (natural calcite crystal of 99.94 wt% purity; the main impurity is Mg at 0.03 wt% according to wet chemical analysis results) and magnesium sulphate (powder, Merck; total metal (other than Mg), nitrogen and chlorine content of 0.05 wt%; main impurity is Ca at 0.04 wt%) in vacuum under ultraviolet irradiation. Figure 1

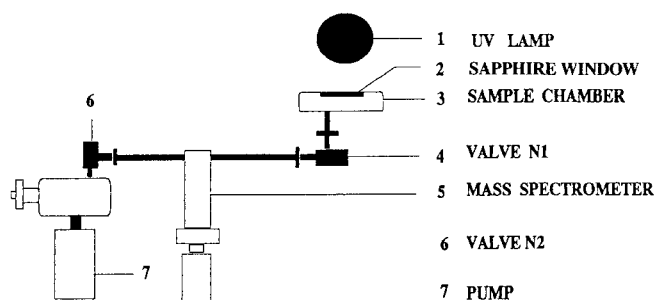


FIG. 1 Experimental set-up for measuring the photo-induced gas release from a solid surface. In the static mode, the vacuum chamber was closed once every hour by the valve N1 and, after 10-min exposure of the sample to ultraviolet radiation, the composition and partial pressure of the gas— CO_2 in the case of CaCO_3 , and SO_2 in the case of MgSO_4 —were measured by a quadrupole mass spectrometer (valve N1 open, N2 closed). In the dynamic mode, both valves N1 and N2 were open, and changes of the partial pressure of the gas species were measured during irradiation.

illustrates schematically our experimental set-up. We made measurements under both static and dynamic conditions (see Fig. 1 legend). The samples were placed into a vacuum chamber with a sapphire window. The total residual pressure was $\sim 10^{-8}$ torr. As the source of ultraviolet radiation we used a 250 W mercury lamp with optical filters in the range 290–580 nm. Mass spectrometry and thermodesorption spectroscopy were used to detect reaction products. The same procedure was used in blank experiments without samples, as well as with a carbonate-free mineral (albite). In both static and dynamic modes it was possible to determine the amount and the rate of gases evolved during irradiation.

Before the decomposition experiments, the samples were carefully degassed at 250–300 °C in high vacuum to remove possible condensed and adsorbed surface contaminants. This cleaning procedure was monitored by mass spectrometry. We considered the sample surface to be clean and ready for measurements if no gases were desorbed during programmed heating of samples in vacuum up to the onset of thermodecomposition. The release of CO_2 from CaCO_3 under ultraviolet radiation at room temperature was clearly established (Figs 2, 3), and the rate of photodecomposition was identical for both modes of operation. No products other than CO_2 were detected, with the exception of CO which appeared after prolonged irradiation (see below). The photo-induced release of CO_2 in blank experiments, as well as

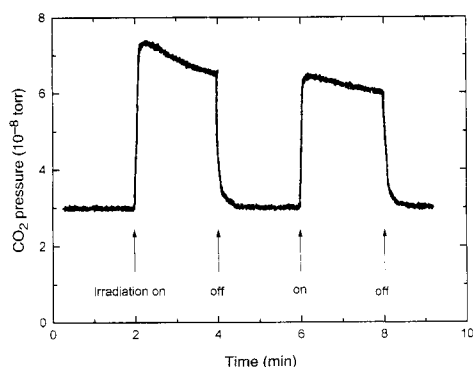


FIG. 2 The response of carbon dioxide partial pressure in the vacuum chamber to a short period of ultraviolet irradiation of the calcite surface at room temperature. The unfiltered radiation of a mercury lamp was used. The data were obtained in the dynamic mode of operation (see Fig. 1 legend).

with an albite sample, was negligible. The photodecomposition rate was highest at the beginning of the irradiation (4×10^{12} molecules $\text{cm}^{-2} \text{s}^{-1}$), and decreased to a near-constant value after several hours of irradiation (2×10^{11} molecules $\text{cm}^{-2} \text{s}^{-1}$). Hence the amount of CO_2 evolved over the course of an experiment corresponds to the decomposition of several monolayers of CaCO_3 . A linear dependence of the initial photodecomposition rate on the light intensity up to fluences of 10^{17} photons $\text{cm}^{-2} \text{s}^{-1}$ was observed.

The spectral dependence of CO_2 release, measured using filtered mercury-lamp radiation, clearly showed a threshold effect at 350–400 nm. This range is far from the edge of the main absorption band of calcite but is coincident with the onset of the very weak additional absorption band with its maximum near 300 nm (ref. 16). The calculated quantum yield of CO_2 release during the initial stage was $\sim 10^{-5}$ molecules per incoming photon at a wavelength near 300 nm.

We showed that the observed photodecomposition of calcite is not related to possible photo-induced desorption of CO_2 molecules (which could be adsorbed on the cleaned sample surface from residual air in the vacuum chamber— $p_{\text{CO}_2} \approx 7 \times 10^{-9}$ torr—during the experiment) as follows; exposure of the sample to a pure CO_2 atmosphere (5×10^{-5} torr, 30 min) before irradiation did not produce measurable changes in the CO_2 photo-induced release. Our thermodesorption studies of the system $\text{CO}_2/\text{calcite}$ did not show a noticeable amount of adsorbed CO_2 under these conditions. The possibility of photo-desorption of some adsorbed CO_2 molecules that remained on the surface even after thermal annealing in vacuum (their coverage can not exceed a monolayer) can be excluded because several monolayers of CO_2 were released under ultraviolet irradiation.

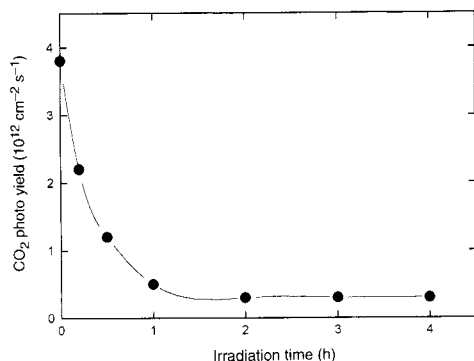


FIG. 3 The release of the carbon dioxide from calcite during prolonged ultraviolet irradiation.

This, together with the absence of reaction products other than CO_2 , also allows us to exclude photocatalytic surface reactions of some unidentified adsorbed molecules as a possible source of CO_2 .

The observed CO_2 release can also not be explained by the heating effect of ultraviolet irradiation because the measured temperature rise of the thermocouple mounted in front of the sample surface was too small to cause the observed rate of CO_2 release. For instance, ultraviolet irradiation in the range 300–330 nm caused the thermocouple temperature to rise up to 50°C and caused a CO_2 pressure rise of 7×10^{-9} torr (in dynamic mode). The same pressure rise in the thermodecomposition experiment was achieved at 160°C . An additional argument against the thermal decomposition of calcite during ultraviolet irradiation was obtained by measuring the intensity ratio of masses 44 (CO_2^+) and 28 (CO^+). During long-term irradiation in the dynamic mode of operation, the value of this ratio decreased from ~ 3 (a value which was the typical ratio for pure CO_2 in our mass spectrometer) to ~ 1 at the end of irradiation. This effect cannot be explained by photodissociation of CO_2 in the gas phase, as it was also observed using filtered irradiation in the range 300–400 nm where decomposition of CO_2 gas molecules does not occur. Moreover, the photodissociation of CO_2 in the gas phase has not been found to lead to our observed ratios of $\text{CO}^+/\text{CO}_2^+$ (refs 17, 18). Apparently the irradiation affects the composition of

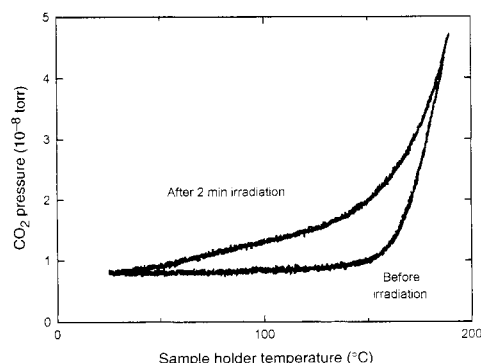


FIG. 4 Thermodecomposition curves for calcite, both before and after a short period of ultraviolet irradiation, obtained under programmed heating of the sample with a mean heating rate of $30^\circ\text{C min}^{-1}$. The difference between the two curves represents the amount of 'stored' carbon dioxide molecules in the damaged surface layer, found to be 7×10^{13} molecules cm^{-2} .

the surface layer so that not only CO_2 , but also CO , was released from this damaged surface layer on ultraviolet irradiation.

Additional evidence of surface damage produced by ultraviolet irradiation was obtained by means of thermodesorption mass spectrometry when we measured the rate of CO_2 release by slow heating during evacuation of the vacuum chamber. As shown in Fig. 4, a clear difference exists between the two thermodesorption curves representing the CO_2 release from irradiated and non-irradiated or annealed samples, indicating the presence of some 'stored' CO_2 in the irradiated surface layer. This additional amount of CO_2 is probably formed during thermostimulated relaxation of the damaged layer with its broken or weakened chemical bonds.

Similar results were observed for photodecomposition of MgSO_4 powder, although the rate of photodecomposition for the sulphate was approximately an order of magnitude lower than for calcite. Another difference between carbonate and sulphate was in the spectral dependences of photodecomposition: the photo-induced release of SO_2 from MgSO_4 took place even in the wavelength range of 500–600 nm, in line with the measured absorption spectra of sulphate samples.

Our results have implications for the surface and atmospheric

chemistry of Mars. The integral flux of photons on the martian surface in the wavelength interval 190–390 nm is $\sim 2 \times 10^{15}$ photons $\text{cm}^{-2} \text{s}^{-1}$ (ref. 19). Taking into account the mean quantum yields for photo-induced release of CO_2 ($\sim 10^{-5}$) due to photodecomposition of carbonates, observed in our study, and assuming a carbonate abundance in martian soil of between $\sim 1\%$ (upper limit)^{13,20} and 0.01% (lower limit)⁶, one obtains fluxes of CO_2 in the near-surface atmosphere layer of $\sim 10^8$ and 10^6 molecules $\text{cm}^{-2} \text{s}^{-1}$. These values are high compared to the estimated CO_2 non-thermal escape rate of $\sim 10^5$ molecules $\text{cm}^{-2} \text{s}^{-1}$ for the martian atmosphere¹², and therefore should be considered in evolutionary models^{13–15}. For martian conditions, relatively high fluxes of photoreleased CO_2 might be maintained by permanent 'refreshing' of the surface by wind-blown high-velocity dust particle bombardment with an erosion rate of $\sim 10^{-3} \text{ cm yr}^{-1}$ (ref. 12). In principle, the photodecomposition of carbonates can, over geological time ($\sim 4 \times 10^9 \text{ yr}$) deliver to the martian atmosphere $\sim 10^{23}$ to 10^{25} molecules cm^{-2} of CO_2 and effect a corresponding transformation of a carbonate layer up to several hundred metres thick. Of course, we estimate here only the intensity of heterogeneous CO_2 photoproduction, and the real intensity will strongly depend on the reverse gas–solid reactions as well as on the shielding of exposed surfaces by aeolian and impact processes. Taking into account high values of the rates of wind-blown particle abrasion²¹ and of dust sedimentation²², we suggest that the thickness of this altered (photoweathered) carbonate-free layer would be the same as that of the dust deposits (up to some metres; ref. 23).

Therefore, our results can also explain why carbonates, although expected to be present on Mars in substantial amounts^{7–9}, have not been detected with certainty by different methods on the martian surface: the surface layer on Mars could

be depleted in carbonates due to photodecomposition. We predict that the attempt to detect carbonates with the Thermal Emission Spectrometer experiment, planned for the 1996 Mars Surveyor Orbiter, will fail; the search for carbonates would need to be conducted in deep layers. □

Received 8 June; accepted 21 November 1995.

- Pollack, J. B. et al. *J. geophys. Res.* **95**, 14595–14627 (1990).
- Calvin, W. M., King, T. V. V. & Clark, R. N. *J. geophys. Res.* **99**, 14659–14675 (1994).
- Bell, J. F., Pollack, J. B., Geballe, T. R., Cruikshank, D. R. & Freedman, R. *Icarus* **111**, 106–123 (1994).
- Blaney, D. L. & McCord, T. B. *J. geophys. Res.* **94**, 10159–10166 (1989).
- Blaney, D. L. & McCord, T. B. *J. geophys. Res.* **100**, 14433–14441 (1995).
- Gooding, J. L., Wentworth, S. J. & Zolensky, M. E. *Meteoritics* **26**, 135–143 (1991).
- Pollack, J. B., Kasting, J. F., Richardson, S. M. & Poliakoff, K. *Icarus* **71**, 203–224 (1987).
- Fanale, F. P., Postawko, S. E., Pollack, J. B., Carr, M. H. & Pepin, R. O. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 1135–1179 (Univ. Arizona Press, Tucson, 1992).
- Gooding, J. L., Arvidson, R. E. & Zolotov, M. Yu. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 626–651 (Univ. Arizona Press, Tucson, 1992).
- Banin, A., Clark, B. C. & Wänke, H. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 594–625 (Univ. Arizona Press, Tucson, 1992).
- Biermann, K. et al. *J. geophys. Res.* **82**, 4641–4658 (1977).
- Luhmann, J. G., Johnson, R. E. & Zhang, M. H. *J. Geophys. Res. Lett.* **19**, 2151–2154 (1992).
- Jakosky, B. M., Pepin, R. O., Johnson, R. E. & Fox, J. L. *Icarus* **111**, 271–288 (1994).
- Pepin, R. O. *Icarus* **111**, 289–304 (1994).
- Haberle, R. M., Tyler, D., McKay, C. P. & Davis, W. L. *Icarus* **109**, 102–120 (1994).
- Weber, M. J. (ed.) *Handbook of Laser Science and Technology* Vol. 4, 2 224–225 (CRC, Boca Raton, Florida, 1986).
- McElroy, M. B. & Donahue, T. M. *Science* **177**, 986–988 (1972).
- Krasnopolsky, V. A. *J. geophys. Res.* **100**, 3263–3276 (1995).
- Mount, G. H. & Rottman, G. J. *J. geophys. Res.* **88**, 5403–5410 (1983).
- Clark, B. C. & Van Hart, D. C. *Icarus* **45**, 370–378 (1981).
- Greeley, R. *J. geophys. Res.* **87**, 10009–10024 (1982).
- Pollack, J. B. et al. *J. geophys. Res.* **84**, 2929–2945 (1979).
- Greeley, R., Lancaster, N., Lee, S., Thomas, P. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 730–766 (Univ. Arizona Press, Tucson, 1992).

ACKNOWLEDGEMENTS. We thank R. Rieder for advice, assistance in conducting experiments and helpful discussions of the results.

Transition from spirals to defect turbulence driven by a convective instability

Q. Ouyang & J.-M. Flesselles

Institut Non Linéaire de Nice, UMR 129 CNRS-UNSA, 1361, route des Lucioles, 06560 Valbonne, France

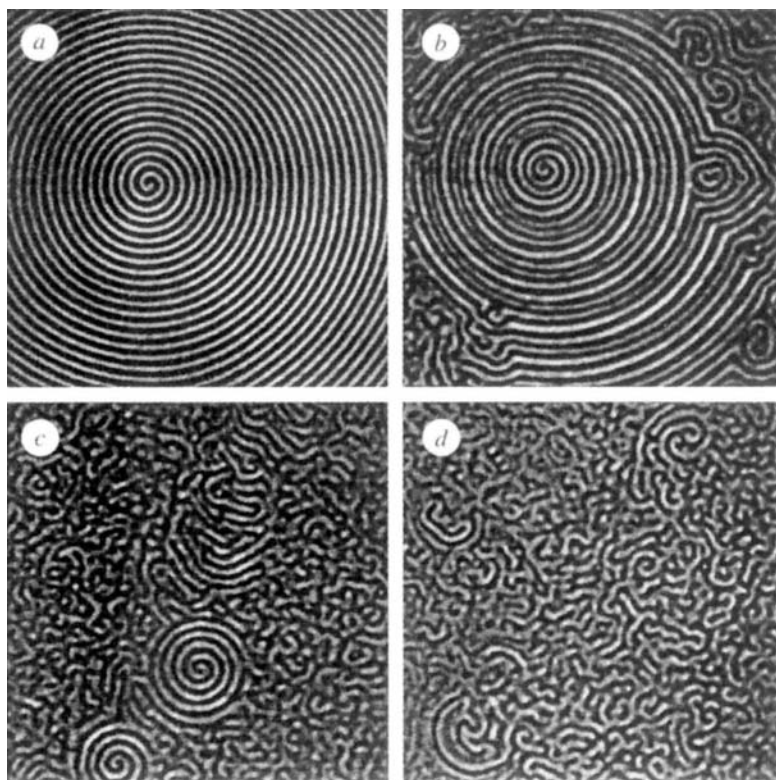
SPATIAL and temporal order often arises in homogeneous systems driven away from equilibrium¹. Sufficiently far from equilibrium, however, highly ordered behaviour typically degenerates into spatiotemporal chaos. The mechanisms underlying this qualitative transition can be clarified by studying a model system comprising a continuous field of coupled nonlinear oscillators. Analyses of the nonlinear partial differential equation that describes this model elucidate generic dynamical features that characterize spatially extended nonequilibrium systems; moreover, they predict an archetypal pathway^{2,3} for the development of spatiotemporal chaos through the spontaneous generation of topological singularities, or 'defects'. The resulting highly irregular state, known as 'defect-mediated turbulence', has not, however, been observed previously in a real system. Here we report an experimental observation of the transition from ordered behaviour to defect-mediated turbulence in the pattern-forming chemical system known as the Belousov–Zhabotinsky reaction. A regular spiral pattern dominates in the ordered state, but as the system is driven further from equilibrium, the spiral becomes unstable and generates hundreds of defects. This observation substantiates a basic mechanism which should underlie the transition to spatiotemporal chaos in a wide variety of pattern-forming systems.

The study of spiral patterns in the Belousov–Zhabotinsky (BZ)

reaction has generated fruitful results^{4–10}. However, except in recent years^{11–13}, all experiments have been conducted in closed systems where the ageing of the chemical solution hinders the study of a steady nonequilibrium state. As a result, spontaneous formation of spiral defects was documented only in the limiting case of lateral instability^{9,10}, close to the end of the reaction. Our experiment was conducted in an open spatial reactor, similar to the one used for the Turing pattern experiments^{14,15}. The BZ reaction occurred in a thin porous glass disk, 0.4 mm thick and 25.4 mm in diameter (Vycor glass, Corning). The porous glass prevented convection in the reaction medium, guaranteeing its reaction–diffusion nature. Each surface of the disk was in contact with a continuously-fed well stirred tank reactor (CSTR) in which the reactant concentrations were kept homogeneous and constant. The components of the BZ reaction were arranged in the CSTRs so that malonic acid resided only in CSTR A and ferroin resided only in CSTR B. The other reactants, sodium bromate and sulphuric acid, were fed with equal concentrations to both CSTRs. A small amount of sodium bromide and sodium dodecyl sulphate were also fed into CSTR A to prevent chemical oscillations and large bubble formation. The chemicals diffused into the porous glass where the BZ reaction occurred and patterns formed. The concentration of sulphuric acid ($[\text{H}_2\text{SO}_4]_0$) was the control parameter.

At low $[\text{H}_2\text{SO}_4]_0$, travelling waves spontaneously emitted by the boundary of the disk undergo a lateral instability¹⁰ and spirals are formed. We then increase $[\text{H}_2\text{SO}_4]_0$ and take advantage of the small light-sensitivity of the reaction to bring one spiral to the reactor centre and to drive the others out off the edge, using a helium–neon laser ($\lambda = 632.8 \text{ nm}$). At low $[\text{H}_2\text{SO}_4]_0$ (0.2–0.5 M), the spiral is a relaxation wave. When $[\text{H}_2\text{SO}_4]_0$ is increased to 0.6 M, the wave becomes sinusoidal (Fig. 1a), which is shown by the single peak of its radial spectrum in Fourier space (Fig. 2a). The rotation period of the spiral is 3.9 s, much faster than a typical spiral observed in closed systems. The spiral wavelength is

FIG. 1 Images *a–d* illustrate the transition from single spiral to spiral defect turbulence as the sulphuric acid concentration was increased. The concentration of sulphuric acid was (in M): *a*, 0.6, stable spiral; *b*, 0.65, onset of convective instability; *c*, 0.7, near absolute instability; *d*, 0.8, full turbulence. Other control parameters were fixed: $[\text{NaBrO}_3]_0 = 0.4 \text{ M}$, $[\text{CH}_2(\text{COOH})_2]_0^A = 0.4 \text{ M}$, $[\text{NaBr}]_0^A = 30 \text{ mM}$, $[\text{SDS}]_0^A = 0.1 \text{ mM}$, $[\text{Ferriin}]_0^B = 1.0 \text{ mM}$, residence time of CSTRs was 18 min, and temperature $23 \pm 0.5^\circ\text{C}$. Here superscript A or B refer to the tanks of the cell; SDS, sodium dodecyl sulphate. The region shown is $12.3 \times 12.3 \text{ mm}$.



0.35 mm, about the thickness of the porous glass. Because of the concentration gradients between the two CSTRs, only a fraction of the glass thickness meets the condition of spiral formation, thus the observed patterns are quasi-two-dimensional.

If one increases $[\text{H}_2\text{SO}_4]_0$ across a critical value (0.63 M), a single spiral becomes unstable: the waves break away from the centre into hundreds of defects, each forming a new spiral, (Fig. 1*b*). After an initial transient response, the spiral–defect interface remains stationary at a fixed reproducible distance from the spiral centre. In Fig. 1*b*, one notices that the pitch of the spiral is not uniform but has local deformations that globally appear as an ‘overspiral’ of pitch 1.2 mm superimposed over the primary one. Figure 3 represents the overspiral extracted from Fig. 1*b*; its long-wavelength character is clear on the radial spectrum of the Fourier transform (Fig. 2*b*). Figure 3*b* shows that the local deformations of the primary spiral are local wavelength modulations, whose amplitude grows with distance from the spiral centre. As $[\text{H}_2\text{SO}_4]_0$ increases, the distance from the spiral-defect interface to the centre of the spiral decreases (Fig. 1*c*). As $[\text{H}_2\text{SO}_4]_0$ crosses another critical value (0.8 M), the stability radius of the spiral reduces to its core size and the system becomes fully turbulent, as shown on Fig. 1*d*. Its spatial Fourier transform shows a broad spectrum (Fig. 2*c*).

Our experiment has been performed in a regime close to the onset of spontaneous sinusoidal oscillations of the BZ reaction. This is the very kind of system that the complex Ginzburg–Landau equation models^{1,16}:

$$\partial A / \partial t = [1 + (1 + i\alpha)\nabla^2 - (1 + i\beta)|A|^2]A \quad (1)$$

This equation describes the time evolution of a two-dimensional pattern, experimentally seen as the real part of $A(x, y, t) \exp(i\omega t)$ where $\omega/2\pi$ is the (fast) frequency of the oscillators and A is a complex number that represents the slow phase and amplitude modulation of the oscillation. α and β are two real numbers that respectively represent the dispersion and the nonlinear frequency shift. To compare with our experiment, the parameters α and β were taken from measurements of Sørensen and Hynne on the BZ reaction¹⁷. As they showed that α relates to the diffusion coefficient of chemicals, and β to the ensemble of reaction kinetics, we assumed that α is a constant and β increases with $[\text{H}_2\text{SO}_4]_0$. The numerical simulation qualitatively agrees with the experiments, as shown by Fig. 4. As β is increased from 0.5 to 0.55, the stable spiral of Fig. 4*a* becomes unstable and generates defects (Fig. 4*b*). Further increase of β decreases the size of stable spirals (Fig. 4*c*). At $\beta = 0.7$, the system becomes fully turbulent (Fig. 4*d*).

A convective instability^{18–20}, similar to the one that drives familiar examples like the movement of cigarette smoke or turbulent jets, is the basic mechanism to explain our experiment. In spatially extended systems, instabilities can either be ‘absolute’ or ‘convective’. This distinction is mainly relevant if there is a natural advection in the system. This advection can be due to a

flow in hydrodynamics, or more generally to travelling waves, as in our experiment. When a single pulse perturbation expands spatially from a given point, it is advected away. In the ‘absolute’ case, the expansion is faster than the advection, so the instability will finally invade the whole space. In the ‘convective’ case, the instability is advected faster than it grows. Any fixed observer, if he experiences the instability at all, will see it passing by.

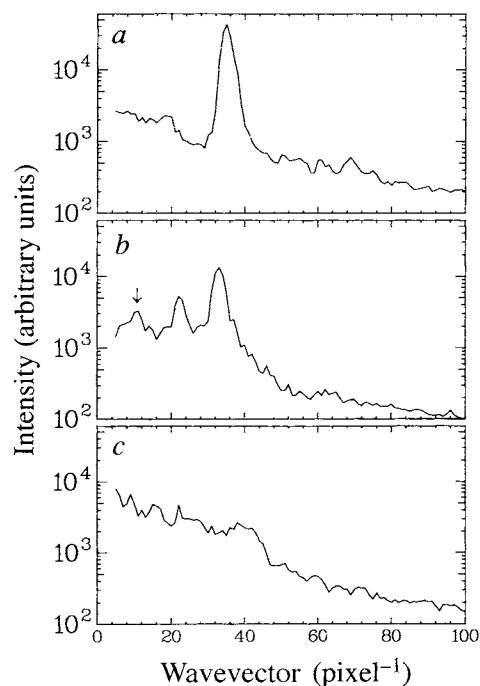


FIG. 2 Power spectra of spatial Fourier transform in the radial direction; panels *a–c* correspond respectively to the states of Fig. 1*a*, *b* and *d*. *a*, Stable spiral; *b*, onset of convective instability; and *c*, full turbulence. The peak at small wavenumbers in *b* (indicated by the arrow) corresponds to the unstable mode in Fig. 1*b*.

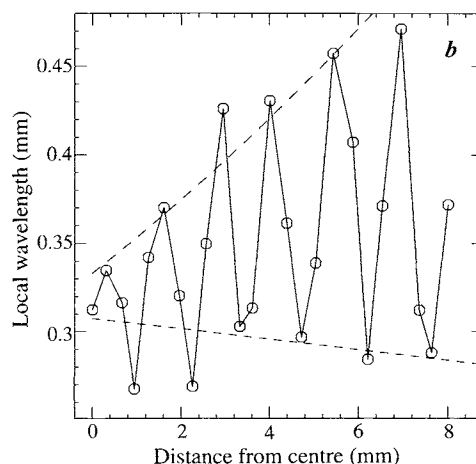
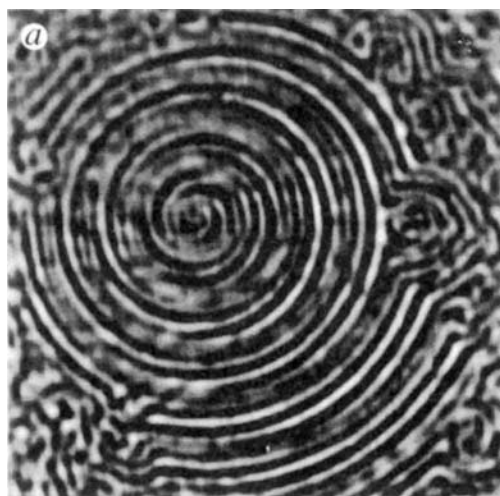


FIG. 3 Features of the wavelength modulation that drives the instability. *a*, The local wavelength modulation of the initial spiral organizes along a spiral with a larger pitch (overspiral), and the same centre as the original spiral. Image obtained by bandpass filtering of Fig. 1*b*; *b*, local wavelength of the pattern along a radial segment taken from Fig. 1*b*. The wavelength of

the primary spiral is modulated around a mean value. The amplitude of the modulation increases with distance from the centre. The period of the modulation is the pitch of the overspiral. Dashed lines are provided to guide the eyes.

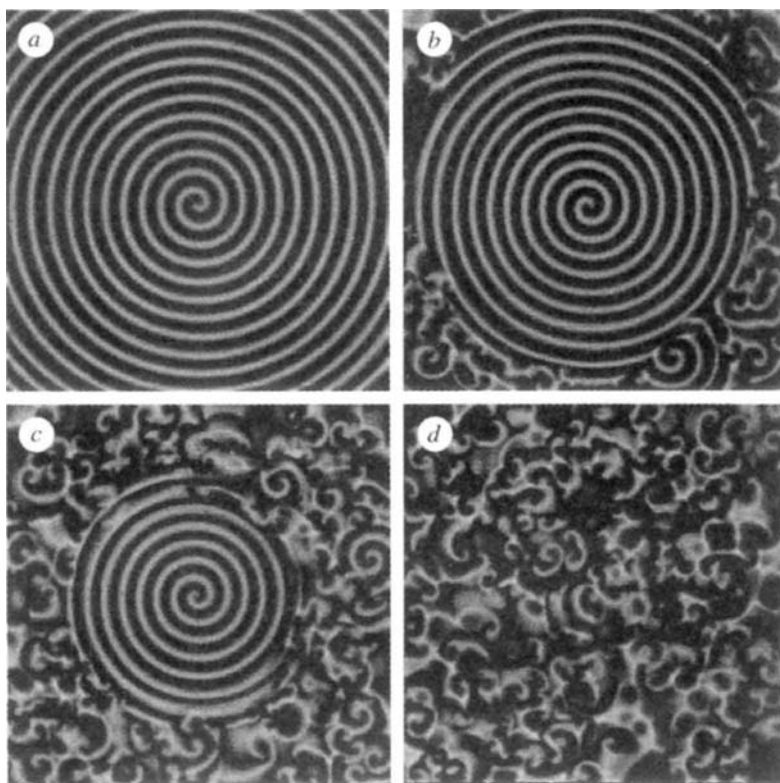


FIG. 4 Numerical integration of the complex Ginzburg–Landau equation (1) generates the same transition to defect turbulence as obtained in the experiments. *a*, $\beta = 0.5$, stable spiral; *b*, $\beta = 0.55$, onset of convective instability. Note the slight wavelength modulation; *c*, $\beta = 0.56$, near absolute instability; and *d*, $\beta = 0.7$, full turbulence. The parameter α was kept fixed at -1.4 . The simulation was conducted in a box of 256×256 grid points with no flux boundary conditions. The grey level of the images is proportional to the real part of A , where A is the solution of the Ginzburg–Landau equation.

In our experiments, as the control parameter ($[\text{H}_2\text{SO}_4]_0$) passes a critical point, the spiral undergoes a wavelength instability (an Eckhaus instability^{1,21,22}) that appears mainly as a periodic modulation of the pitch. This instability nucleates at the centre of the spiral, where it has a vanishing amplitude, and grows as it is advected away by the spiral waves, at the spiral's group velocity³. Thus at some distance from the spiral centre, we observe a stable spiral and its long-wavelength modulation, (Figs 1*b*, 4*b*). Far enough from the centre, the instability has grown so much that it destroys the carrying waves. This is a well known theoretical model of topological defect nucleation that leads to defect-mediated turbulence². According to this picture, the stability radius of the spiral should be proportional to the group velocity, and inversely

proportional to the growth rate of the instability. If we assume that the growth rate increases as the control parameter moves away from the onset of the instability, it should increase with $[\text{H}_2\text{SO}_4]_0$, and the size of the stable spiral should shrink (Figs 1*c*, 4*c*), in qualitative agreement with experimental observations. If $[\text{H}_2\text{SO}_4]_0$ is above a certain value, the group velocity can no longer balance the growth rate; the regime of absolute instability is attained, corresponding to the fully turbulent state (Figs 1*d*, 4*d*). □

Received 6 July; accepted 28 November 1995.

1. Cross, M. C. & Hohenberg, P. C. *Rev. mod. Phys.* **65**, 851–1112 (1993).
2. Couillet, P., Gil, L. & Lega, J. *Phys. Rev. Lett.* **62**, 1619–1622 (1989).
3. Weber, A., Kramer, L., Aranson, I. S. & Aranson, L. *Physica D* **61**, 279–283 (1992).

4. Zaikin, A. N. & Zhabotinsky, A. M. *Nature* **225**, 535–537 (1970).
5. Winfree, A. T. *Science* **175**, 634–636 (1972).
6. Agladze, K. I., Krinsky, V. I. & Pertsov, A. M. *Nature* **308**, 834–835 (1984).
7. Maselko, J. & Showalter, K. *Physica* **D49**, 21–32 (1991).
8. Steinbock, O., Zykov, V. & Müller, S. C. *Nature* **366**, 322–324 (1993).
9. Markus, M., Kloss, G. & Kusch, I. *Nature* **371**, 402–404 (1994).
10. Nagy-Ungvári, Zs. & Müller, S. C. *Int. J. Bifurc. Chaos* **4**, 1257–1264 (1994).
11. Noszticzius, Z., Horsthemke, W., McCormick, W. D., Swinney, H. L. & Tam, W. Y. *Nature* **329**, 619–620 (1987).
12. Skinner, G. & Swinney, H. L. *Physica* **D48**, 1–16 (1991).
13. Dulos, E., Boissonade, J. & De Kepper, P. *Physica* **A188**, 120–131 (1992).
14. Oyang, Q. & Swinney, H. L. *Nature* **352**, 610–612 (1991).
15. Ouyang, Q. & Swinney, H. L. *Chaos* **1**, 411–420 (1991).
16. Kramer, L., Hynne, F., Sørensen, P. G. & Walgraef, D. *Chaos* **4**, 443–452 (1994).

17. Sørensen, P. G. & Hynne, F. J. *phys. Chem.* **93**, 5467–5474 (1989).
18. Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* (Pergamon, Oxford, 1959).
19. Huerre, P. in *Propagation Far from Equilibrium* (ed. Weisfred, J. E., Brand, H. R., Manneville, P., Albinet, G. & Boccaro, N.) 340–353 (Springer, Berlin, 1988).
20. Manneville, P. in *Structures Dissipatives, Chaos et Turbulence* 349–352 (Aléa-Saclay, 1991).
21. Fauve, S. in *Instabilities and Nonequilibrium Structures* (eds Tirapegui, E. & Villareal, D.) 63–88 (Reidel, Dordrecht, 1987).
22. Janiaud, B., Pumi, A., Bensimon, D., Croquette, V., Richter, H. & Kramer, L. *Physica* **D55**, 269–286 (1992).

ACKNOWLEDGEMENTS. We thank Ch. Dupont for doing the numerical simulation, P. Coulet for suggesting the theoretical interpretation, and A. Belmonte, L. Gil, V. Krinsky, J. Lega and V. Voignier for discussions. Q.O. is on leave from the Physics Department, University of Texas at Austin.

Pathways of water between the Pacific and Indian oceans in the Indonesian seas

Arnold L. Gordon* & Rana A. Fine†

* Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964-8000, USA

† Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida 33149, USA

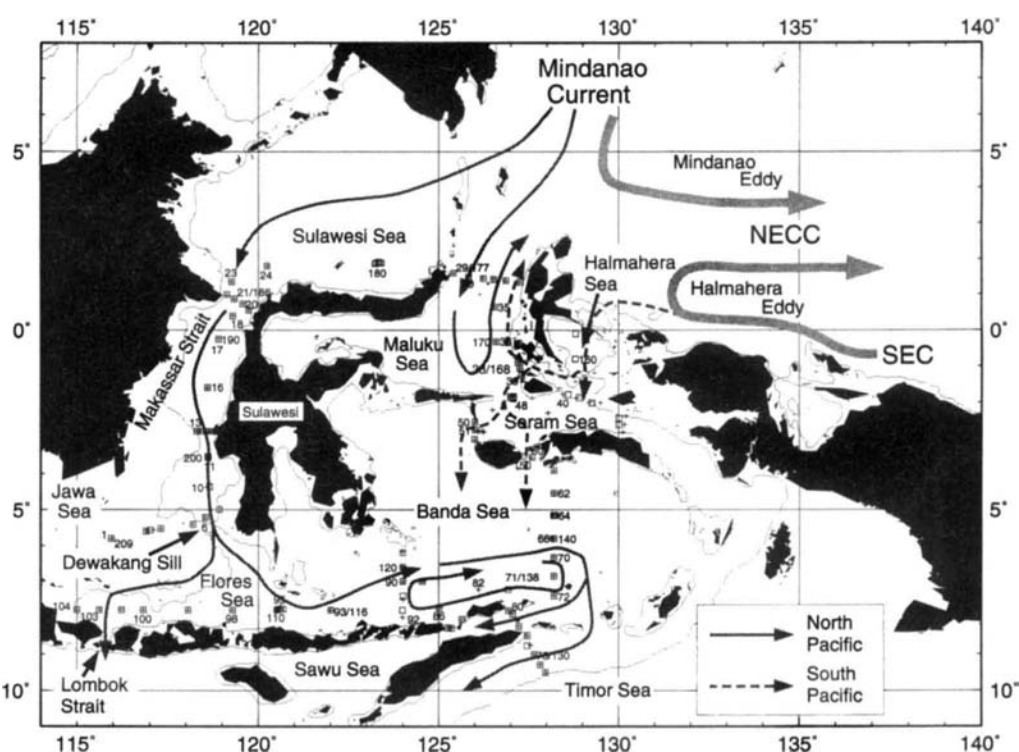
THE physical structure of the Pacific and Indian oceans is substantially affected by the inter-ocean transport of excess fresh water from the North Pacific Ocean through the Indonesian seas^{1,2}. The efficiency of this transport is an important regulator of the meridional overturning of these oceans^{1,2}, and hence perhaps of the global thermohaline circulation³; in addition, the seepage of warm water out of the Pacific affects the volume of the western Pacific warm pool, and thus may influence El Niño events²⁴. But the sources, pathways and physical properties of the

Indonesian throughflow are not well enough characterized to allow its influence on ocean circulation and the climate system to be quantified. Here we report salinity, temperature and chemical-tracer data from the Indonesian seas which show that the throughflow is dominated by two components: one of low-salinity, well ventilated North Pacific water through the upper thermocline of the Makassar Strait, and the other of more saline South Pacific water through the lower thermocline of the eastern Indonesian seas. Seasonal (monosonal) variations in the ratio of these components, perhaps modulated by El Niño conditions, imply the existence of potentially important variable feedbacks to the ocean circulation and climate system.

Recent observational studies offer a wide range of estimates for the Indonesian throughflow (Fig. 1), from 7 to $18.6 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ (refs 4–6) including $1.7 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ within the Lombok Strait⁷. Density-driven overflow via the 1,940-m-deep Lifamatola Passage into the depths of the Banda Sea⁸ contributes a minor amount to the total throughflow, approximately $10^6 \text{ m}^3 \text{ s}^{-1}$ (1 sverdrup). Water mass indicators suggest^{3,9–12} that the throughflow occurs mainly within the thermocline and is drawn from the North Pacific, although this too remains a point of debate^{13,14}.

The Arlindo programme¹³ measured the thermohaline and chlorofluorocarbon (CFC-11 and CFC-12) stratification within

FIG. 1 Map of the oceanographic stations where measurements were obtained by the Arlindo programme, place names referred to in the text, and a schematic representation of the water spreading patterns in the Indonesian seas. The Arlindo oceanographic observations within the Indonesian seas consist of measurement of the temperature, salinity and oxygen stratification (CTD stations) and chlorofluorocarbon (CFC-11 and CFC-12) determinations. The data were obtained from the Indonesian R/V *Baruna Jaya I* during the southeast monsoon of 1993 (6 August to 12 September, stations 1–103, + symbols) and northwest monsoon of 1994 (25 January to 3 March, stations 104 to 209, □ symbols, many co-located with 1993 stations). The positions of stations used in Figs 2 and 3 are indicated. A schematic of the thermocline circulation pattern, based on water mass analysis, is indicated by the arrows denoting spreading pathways of North and South Pacific thermocline water masses. NECC is the North Equatorial Countercurrent, SEC is the South Equatorial Current. The Makassar Strait, carrying North Pacific water, is the dominant throughflow pathway. There appears to be a minor contribution of North Pacific water along pathways east of Sulawesi. The South Pacific supplies relatively saline lower-thermocline water and all of the deeper water via the



pathways east of Sulawesi. The lack of evidence for advective spreading indicates that there is not much transport associated with the South Pacific thermocline water.

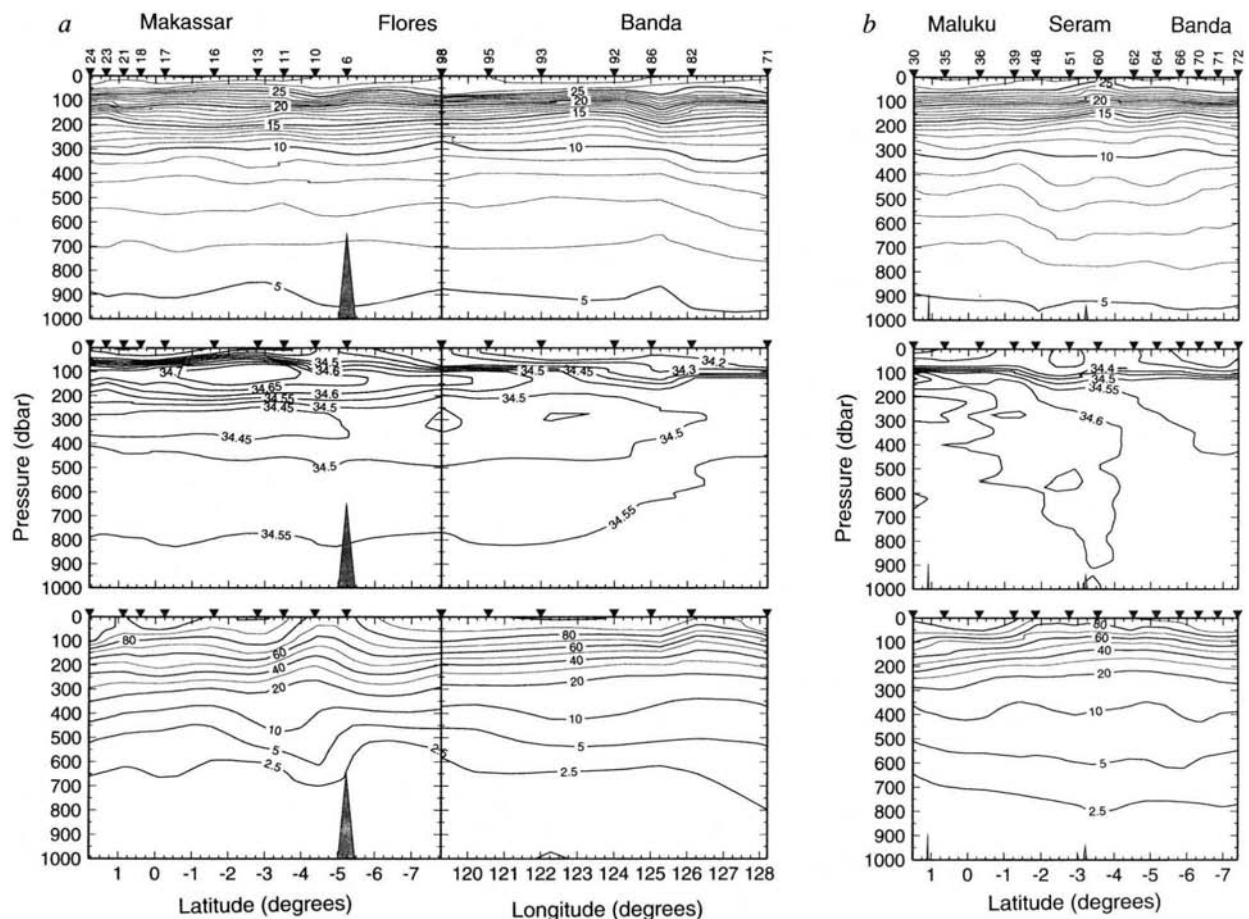


FIG. 2 Sections of (top to bottom in *a* and *b*) potential temperature ($^{\circ}\text{C}$), salinity (p.p.t.) and percentage saturation (relative to the present atmosphere) of CFC-11 during the 1993 southeast Monsoon. Filled triangles at the top of each panel represent the station positions. The triangle on the bottoms of each panel at station 6 marks the shallowing at the Dewakang sill (Fig. 1). *a*, Sections along the Makassar Strait, Flores Sea into the southern Banda Sea. The stratification along the Makassar Strait and Flores Sea reveals the primary thermocline throughflow water masses of the region: the salinity maximum (S-max) near 125 dbar (1 dbar pressure $\approx 1\text{ m}$) and mode water (marked by a slight reduction in the temperature gradient near 17°C , 175 dbar at the same stations displaying a strong

S-max) are derived from the North Pacific subtropical region; the salinity minimum (S-min) near 300 dbar is derived from the North Pacific Intermediate Water. *b*, Sections across the Maluku Sea, Seram Sea and Banda Sea. The northern Maluku Sea thermocline has similar properties to that of the Makassar Strait, but in the southern Maluku and in the Seram Sea the North Pacific S-max and S-min cores are absent, replaced by a lower-thermocline salinity maximum derived from the South Pacific thermocline. The southern Banda Sea thermocline more closely resembles that of the Flores, indicating it as its source. The corresponding NWM section (not shown) reveals slightly more concentrated South Pacific S-max in the Seram Sea.

the Indonesian seas during the 1993 southeast monsoon (SEM) and 1994 northwest monsoon (NWM) (Fig. 1). North Pacific salinity and CFC stratification features penetrate into the Banda Sea within the westernmost deep pathway from the Makassar Strait through the Flores Sea (Fig. 2*a*). Within the eastern throughflow route of Maluku to Seram to the Banda Sea (Fig. 2*b*), the North Pacific stratification features are discontinuous. Although the salinity maximum (S-max) and minimum (S-min) layers are persistent along the Makassar route, they do attenuate with distance from their entry at the Sulawesi Sea, as vertical mixing and dilution with local freshwater input diminishes the salinity stratification¹⁰. The S-max decrease is not uniform, as local pockets of elevated S-max are observed in the Makassar Strait in both monsoons, possibly denoting an uneven injection of North Pacific water.

In the southern Maluku and Seram seas (Fig. 2*b*), the North Pacific middle and lower thermocline is replaced by water too salty to be derived from the North Pacific. This water is South Pacific thermocline water which enters the Seram Sea through the Halmahera Sea, drawn from the New Guinea Coastal Undercurrent¹¹. Similarity of the Maluku and Seram Sea low-salinity

water within the upper 100 dbar to that of the Banda Sea (with dissimilarity to Sulawesi Sea surface water) suggests that the surface flow is towards the north, consistent with surface current charts¹⁵.

CFC-11 can be used to identify the water masses in the Indonesian seas that are involved in decadal timescale climate processes. Based on the CFC-11 partial pressures¹⁶, waters in the upper 200 dbar have been in contact with the atmosphere in the past 5 to (at most) 18 years in Makassar as compared with 13 to (at most) 22 years in Banda. The strongest age gradient occurs above 300 dbar (Fig. 2), similar to the Mindanao Current¹¹. The recently ventilated (with respect to atmospheric gases) waters from the Mindanao Current observed in the Makassar above 300 dbar, particularly during the SEM, are mixed down and redistributed. In addition, during the NWM there is an input of South Pacific lower thermocline and intermediate water that is more highly saturated with CFC-11. As a consequence these layers in the Banda Sea waters are more highly saturated in CFC-11 than those of the Makassar or Maluku (Fig. 3), and they ultimately spread into and ventilate the Indian Ocean.

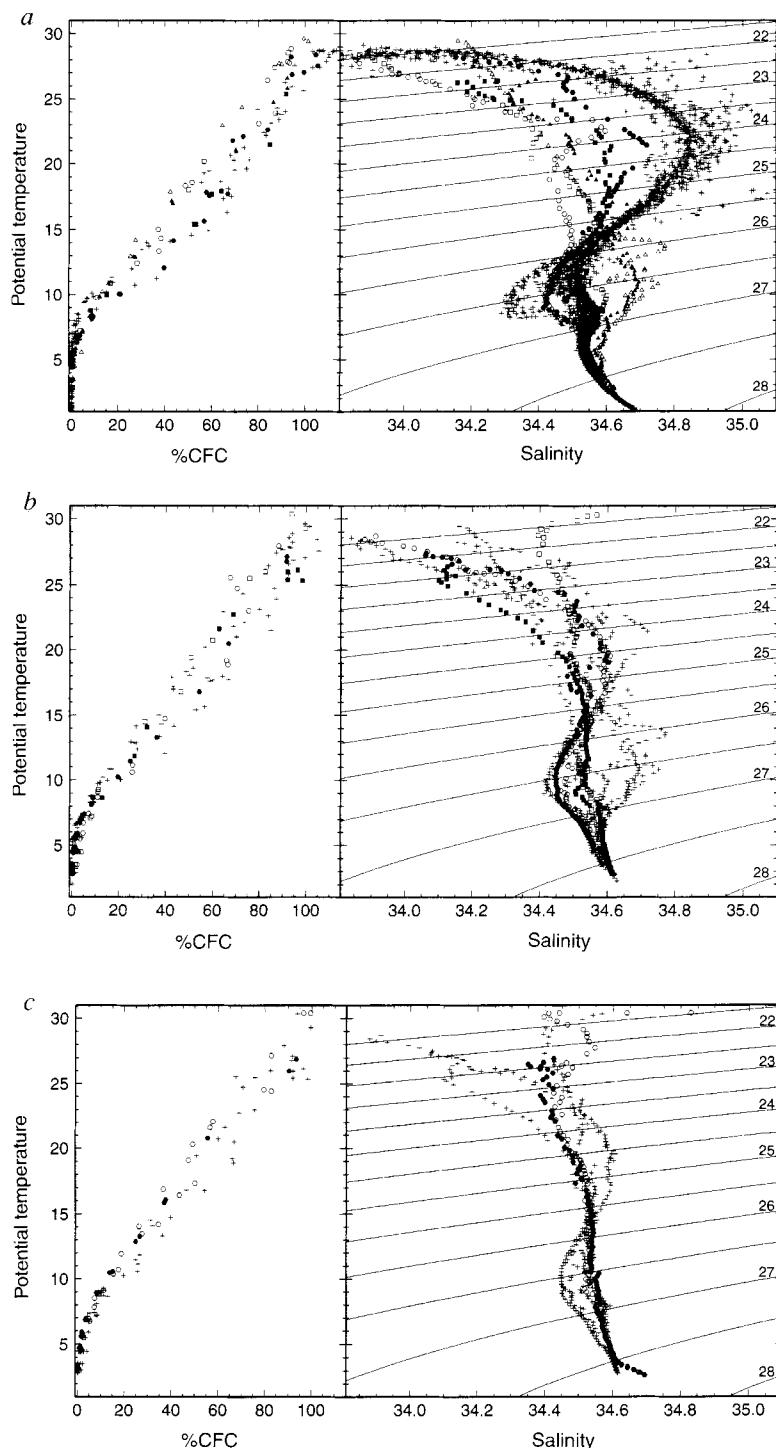
In the northern entrance to the Makassar Strait and the Maluku

Sea, the relationships between potential temperature T and salinity (T/S) or percentage saturation with CFC-11 ($T/\%CFC$) (Fig. 3a) are very similar—though with diminished S -max, S -min and CFC-11 cores—to that of the Mindanao Current¹³. At the S -min core, the Makassar T/S relationship bears a closer resemblance to the inshore edge of the Mindanao Current, whereas the Maluku T/S is similar to the off-shore Mindanao Current. We suspect that the fine structure in the Maluku S -max layer is derived from the saltier South Pacific thermocline water observed in the southern Maluku Sea and the Seram Sea; this is particularly evident in the northeastern edge of the Maluku, suggesting an anti-clockwise circulation pattern within the Maluku thermocline (Fig. 1). The South Pacific thermocline water is seen in more concentrated form in the Seram Sea, and can only be derived from

the Halmahera Sea. We note that the continuity of the North Pacific thermocline water between the Maluku Sea and Banda Sea is severed in the Seram Sea by the South Pacific thermocline water. However, deeper South Pacific water (below the Halmahera Sea sill depth of 500 dbar, and which therefore must be drawn from the Mindanao Current) is continuous from the Maluku to the Banda Sea, spilling into the deep Banda Sea over the Lifamatola sill¹⁸.

In the Makassar, the S -max and $\%CFC$ are decidedly weaker in the NWM. We suggest that this reflects a relaxation of the throughflow during the NWM, with continued mixing in the Makassar Strait acting to attenuate the resident S -max and CFCs. The prevalence of 50-dbar-thick homogeneous layers or 'steps' in the Makassar thermocline indicate the presence of strong mixing¹⁷. Relaxation of the throughflow in the NWM is

FIG. 3 Scatter plots of potential temperature/salinity (T/S) and potential-temperature/percentage saturation CFC-11 ($T/\%CFC$) relationship for selected stations within the Indonesian Seas. The curved lines within the T/S plot marked 22–28 are lines of equal density (isopycnals) in σ_θ nomenclature. Due to the sparser resolution for CFCs, multiple and nearby stations are used to fill out the $\%CFC$ profile. Solid symbols represent the southeast Monsoon (SEM); open symbols represent the northwest Monsoon (NWM). a, The northern Indonesian seas. Circles show Makassar Strait stations, the T/S data are derived from stations 21 and 186; the $\%CFC$ data are derived from the stations 17, 18, 21, 186, 189 and 190 for the NWM. Boxes show the Maluku Sea, the T/S data are derived from stations 29 and 177; the $\%CFC$ data are derived from the stations 30 and 176. Triangles show the Seram Sea, the T/S data are derived from stations 38 and 168; the $\%CFC$ data are derived from the stations 38, 39, 167 and 169. *R/V Moana Wave* stations 3–36 (stations from May 1989 spanning the Mindanao Current west of 140° E between 8° and 9° N; ref 23) are shown as a +, as reference for the Arindo T/S and $T/\%CFC$ scatter. The *Moana Wave* stations west of 127.67° E (within 115 km of the continental shelf) are responsible for the low-salinity feature in the 8 – 13° C temperature interval, making a rather abrupt transition from more saline water to the east. It is this water that can be traced into the Makassar Strait. b, The southern Indonesian seas. Circles show the Flores stations, the T/S data are derived from stations 93 and 116; the $\%CFC$ data are derived from the stations 96, 97, 113 and 114. Boxes show the Banda Sea, the T/S data are derived from stations 71 and 138; the $\%CFC$ data are derived from stations 70, 71, 138 and 139. The northern data (from Fig. 3a) are shown as a + for reference. c, Timor Sea. Circles show the T/S and $T/\%CFC$ relation for the Timor Sea, the T/S data are derived from stations 75 and 130; the $\%CFC$ data are derived from the stations 74, 75, 128, 129, 130. The Flores and Banda sea (Fig. 3b) are shown as a + for reference.



consistent with the model results¹⁸. The Flores Sea *S*-max and %CFC are stronger in the NWM, suggesting a monsoon phase lag of the *S*-max core layer between the Makassar and Flores. Considering the path length and width of Makassar Strait, a mean throughflow speed for the upper 300 m of 7 cm s^{-1} (corresponding to a transport of 5 sverdrup) would account for the phase lag.

In the Flores and Banda seas (Fig. 3b), the properties are similar to that of Makassar Strait, though in the Banda Sea the lower thermocline has somewhat elevated salinity indicating encroachment of South Pacific water via the Seram Sea¹⁹ and the accumulated effects of mixing as discussed above. The Flores and Banda *T/S* curves match below the 5°C isotherm. As this water is deeper than the Dewakang sill of southern Makassar Strait, it is totally derived from the Seram Sea. Above the 4°C isotherm, the Timor Sea properties match that of the Banda Sea (Fig. 3c), indicating that the Banda Sea is the source of Timor Sea stratification. Below 4°C the Timor water column is blocked from the southern Banda Sea by a 1,400-m sill⁴; the deeper more saline water within the Timor Sea is derived from the Indian Ocean⁵.

The Arlindo data set unequivocally shows that the throughflow source and path is that of North Pacific thermocline water flowing from the Makassar Strait into the Flores and Banda Seas before curling southwards into the Timor Sea and Indian Ocean. East of Sulawesi there is little evidence of North Pacific throughflow within the upper thermocline, but the presence of relatively salty, lower-thermocline water suggests a direct contribution from South Pacific Ocean via the Halmahera Sea. Below the thermocline and Halmahera sill depth, the source of the Indonesian water is drawn from South Pacific but by a more indirect route involving transfer of water from the Mindanao Current to the Maluku Sea.

The inter-ocean throughflow pathway appears to follow the westernmost course of least impedance. To model realistically the throughflow pathway and source, it is becoming increasingly clear that it is critical to model very accurately the vigorous vertical mixing and geometry of the Indonesian region. This is needed not only to derive a realistic transport pattern, but also to reflect the correct ratio of North to South Pacific contribution to the throughflow at various depths²⁰. Results of a recent modelling study²¹, although fairly realistic, force most of the Makassar throughflow through the Lombok Strait rather than into the Flores Sea and fail to show saline South Pacific lower-thermocline water infiltrating the Seram Sea, instead showing surface South Pacific water entering the Seram. A model that fails to faithfully reproduce the source water ratio would be forced to find unrealistic ways to balance the hydrological cycle of the Pacific and Indian oceans²².

The Arlindo programme occurred during a prolonged El Niño/Southern Oscillation (ENSO) event, though there was some relaxation of the southern oscillation index in January 1994, before plunging again into an El Niño condition in early 1994. The seasonal cycle of the source water, as revealed by the Arlindo data, opens the possibility of that the ratio of the North to South Pacific throughflow changes in response to the phase of ENSO, which may in turn provide feedback to ENSO by influencing the extent of the warm pool. The El Niño phase with diminished warm-pool presence in the western Pacific may be an analogue to the northern-winter, NWM condition, when the warm pool shifts southwards²⁴, diminishing its presence at the throughflow entrances. The La Niña phase would then correspond to the northern summer, SEM condition, when the warm-pool presence at the entrances is strengthened²⁴. □

6. Meyers, G., Bailey, R. & Worby, A. *Deep-Sea Res. Pt 1* **42**, 1163–1174 (1995).
7. Murray, S. & Arief, D. *Nature* **333**, 444–447 (1988).
8. Van Aken, H. M., Punjawan, J. & Saimima, S. *Neth. J. Sea Res* **22**, 315–339 (1988).
9. Fine, R. *Nature* **315**, 478–480 (1985).
10. Ffield, A. & Gordon, A. *J. phys. Oceanogr.* **22**, 184–195 (1992).
11. Fine, R., Lukas, R., Bingham, F., Warner, M. & Gammon, R. *J. geophys. Res.* **99**, 25063–25080 (1994).
12. Wyrki, K. *J. geophys. Res.* **92**, 12941–12946 (1987).
13. Gordon, A. *J. Phys. Oceanogr.* **25**, 1560–1567 (1995).
14. Godfrey, S., Hirst, A. & Wilkin, J. *J. Phys. Oceanogr.* **23**, 1087–1098 (1993).
15. Wyrki, K. NAGA Rep. 2 (Scripps Inst. Oceanogr, La Jolla, 1961).
16. Doney, S. C. & Bullister, J. L. *Deep-Sea Res. Pt 1* **39**, 1857–1883 (1992).
17. Ffield, A. thesis, Columbia Univ. (1994).
18. Kindle, J. C., Hurlburt, H. E. & Metzger, J. *Proc. Western Pacific Int. Mtg & Workshop on TOGA COARE* (eds Picaut, J., Lukas, R. & Delcroix, T.) 355–365 (Univ. Hawaii, Honolulu, 1989).
19. Gordon, A., Ffield, A. & Ilahude, A. G. *J. geophys. Res.* **99**, 18235–18242 (1994).
20. Nof, D. *J. phys. Oceanogr.* **25**, 1369–1383 (1995).
21. Miyama, T., Awaji, T., Akitomo, K. & Imasato, N. *J. Geophys. Res.* **100**, 20517–20541 (1995).
22. Piola, A. & Gordon, A. *J. phys. Oceanogr.* **14**, 747–753 (1984).
23. Johnson, G. & Toole, J. *Deep-Sea Res.* **40**, 371–394 (1993).
24. Wyrki, K. in *Proc. Western Pacific Int. Mtg & Workshop on TOGA COARE* (eds Picaut, J., Lukas, R. & Delcroix, T.) 99–109 (Univ. Hawaii, Honolulu, 1989).

ACKNOWLEDGEMENTS. We thank M. T. Zen, I. Soesilo, A. Soegiarto, K. Romimohtarto and A. Nontji for their help, and A. G. Ilahude for assistance in organizing the Arlindo cruises. We also thank the captain and crew of the *R/V Baruna Jaya I*, and the US and Indonesian scientific team, for their fine support, particularly B. Huber, P. Mele, A. Ffield, K. Sullivan, K. Maillet, J. Belinne and Mushwerry.

Thermal structure of a fossil mantle diapir inferred from the distribution of mafic cumulates

Georges Ceuleneer, Marc Monnereau & Isma Amri

Centre National de la Recherche Scientifique, UPR 234, Observatoire Midi-Pyrénées, 14, avenue Edouard Belin, 31400 Toulouse, France

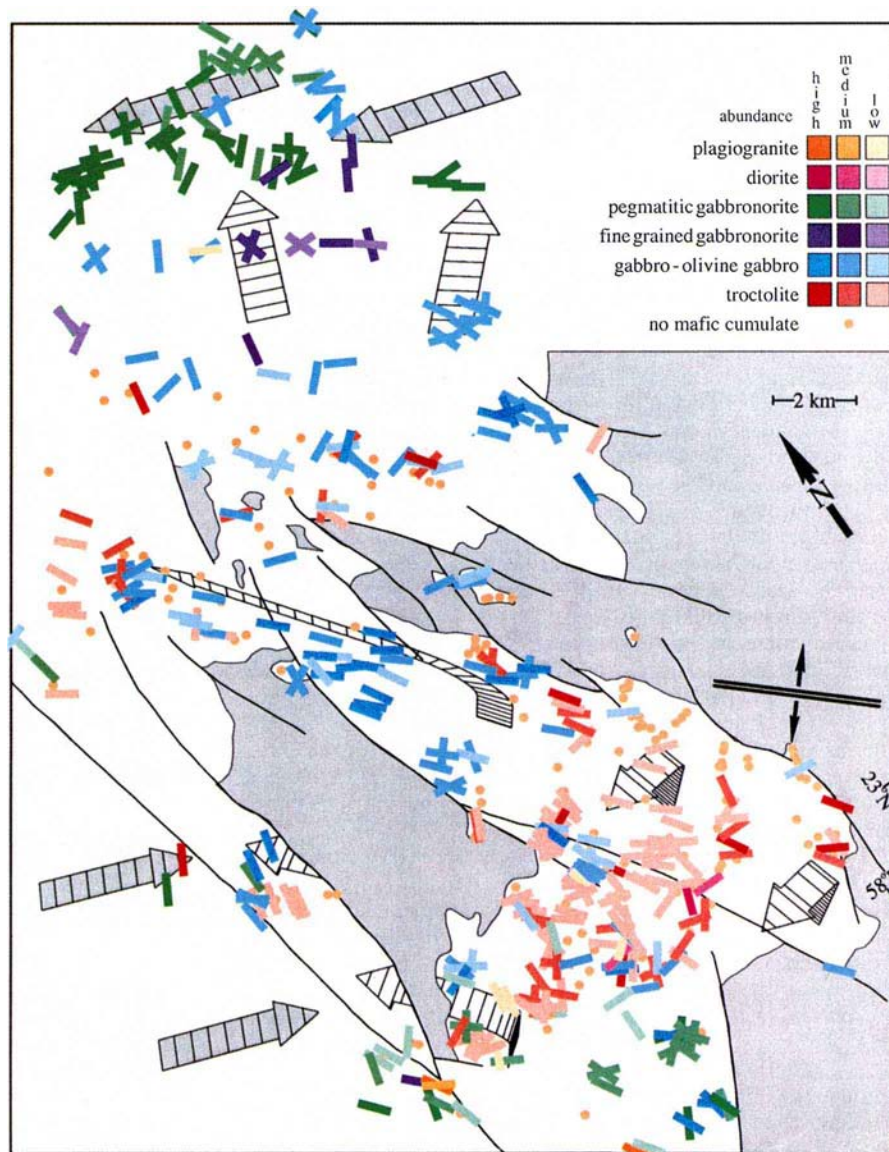
MAFIC lithologies constitute a minor but ubiquitous component of the mantle section of ophiolites; they are currently viewed as cumulates left by basaltic melts travelling from the melting region to the surface. These features may be used to constrain the mechanisms of melt migration in the mantle, provided their geophysical context of emplacement can be established in some detail. Here we report on the nature and distribution of mafic cumulates in the harzburgites of the Oman ophiolite, within and around a mantle diapir frozen beneath a spreading axis^{1–6}. We show that their composition, texture and field characteristics are systematically related and display a concentric zoning centred on the diapir. We propose that this zoning reflects the thermal structure of the diapir and of the associated melt plumbing system. Assuming a parental liquid with the composition of mid-ocean-ridge basalt^{7–9}, the cumulate compositions show that the melt temperature exceeded $1,230^\circ\text{C}$ in the innermost part of the diapir and decreased progressively down to $<1,100^\circ\text{C}$ in the surrounding lithospheric mantle. The transition between channelled porous flow and dyking is observed in an outer shell of the diapir, where the melt crystallized along the plagioclase–olivine cotectic (temperature ranging from $1,180$ to $1,210^\circ\text{C}$).

Mafic cumulates scattered within outcrops of mantle peridotites display a wide spectrum of field characteristics, indicative of the diversity of small-scale processes involved in melt migration: convincing evidence has been reported for contrasting mechanisms such as fracture, dissolution and compaction^{4,5,10–15}. There are, however, important differences among authors concerning the attribution of a mechanism (or combination of mechanisms) to a particular geophysical context. Among others, the problem of the transition between intergranular porous flow and dyking—the ‘brittle–ductile’ transition for melt transport in the mantle—is still eagerly debated. Do fractures develop in the partially molten mantle or are they restricted to the cooler litho-

Received 26 May; accepted 28 November 1995.

1. Piola, A. & Gordon, A. *J. phys. Oceanogr.* **16**, 2184–2190 (1986).
2. MacDonald, A. *J. geophys. Res.* **98**, 6851–6868 (1993).
3. Gordon, A. *J. geophys. Res.* **91**, 5037–5047 (1986).
4. Cresswell, G., Frische, A., Peterson, J. & Quadfasel, D. *J. geophys. Res.* **98**, 14379–14390 (1993).
5. Fleux, M. et al. *Deep-Sea Res. Pt 1* **41**, 1091–1130 (1994).

FIG. 1 Distribution of mafic cumulates in the harzburgites of the Maqсад area. Abundance refers to the proportion of cumulates relative to the residual harzburgites, averaged over a surface of a few hundred square metres. Low, medium and high abundances correspond roughly to < 1%, 1–5% and > 5% of cumulates respectively. Millimetre-sized veins have been excluded from this map for clarity. The crustal section is shown in grey and includes, in this map, the dunites of the transition zone between the harzburgites and layered gabbros. Large arrows show schematically the mantle flow pattern recorded by the harzburgites in the Maqсад area^{1–6}. As the massif is sub-horizontal (the Moho discontinuity presents a tilt of 5–10° towards the southeast), the mantle structures can be directly interpreted in terms of planform mantle flow. The central part of the mapped area is characterized by a zone of vertical plastic flow—the Maqсад diapir—narrowing abruptly toward the northwest. Only the southwestern flank of the Maqсад diapir is exposed. To the north-east, it probably continues below the crustal section¹⁶. All around it, the mantle flow displays a diverging pattern that allows the accurate location of the palaeo-spreading axis related to the Maqсад diapir (black double line). The precise orientation of the Maqсад ridge is deduced from the azimuth of the sheeted-dyke complex at the top of the crustal section (Fig. 3). The zone of diverging flow (non-shaded arrows) is referred to as the periphery of the diapir. The Maqсад spreading structure opened in a lithospheric lid formerly accreted along a north–south ridge⁴. East–west-trending mantle lineations related to this north–south ridge are observed in the north-eastern and southwestern parts of the mapped area (shaded grey arrows).



sphere? The precise tectonic setting and mantle temperature at the time when cumulate features were emplaced is a critical parameter for answering this question: an outcrop of peridotites may integrate a multi-stage structural and/or magmatic history that is generally quite difficult to decipher. In that respect, the Maqсад area in the Oman ophiolite presents a particularly readable structural record that should make the physical interpretation of cumulate features easier than elsewhere. The mantle and crustal structures in the Maqсад area can be attributed to a spreading event along a N130° E-striking ridge^{4,16,17}. Shear-sense polarity in the peridotites allows the accurate location of the Maqсад palaeo-ridge axis⁴; vertical plastic flow lines allow the delineation of the associated mantle upwelling (the Maqсад diapir)^{1–6}. Before it ceased activity, the Maqсад ridge underwent 25 km of spreading and the Maqсад diapir made its way through the surrounding lithosphere to the base of the crust⁴.

Mafic cumulates in the Maqсад harzburgites belong to several lithological groups that are not randomly distributed relative to the plastic flow structures (Fig. 1). The diapir itself (the zone of vertical mantle flow lines) is characterized by an exceptionally low abundance of cumulates. The largest area virtually devoid of any mafic cumulates forms a 1-km-wide corridor ~6 km long parallel to the Maqсад ridge in the central part of the diapir. These 'barren' harzburgites contain, however, isoclinally folded dunitic

layers locally associated with concordant chromite ore bodies. Dunite layers in Oman harzburgites have been shown to form as a result of incongruent dissolution of enstatite in response to migration of basaltic melt with the geochemical character of mid-ocean-ridge basalt (MORB)¹⁴. Chromite could be the very first cumulate phase that precipitated from this melt, although a more complex crystallization history of chromite cannot be ruled out¹⁸.

The rare mafic cumulates present in the diapir are troctolitic in composition. The contact relationships of these troctolites with their host harzburgites range from diffuse to sharp. Diffuse troctolites occur as centimetre- to metre-thick segregations, typically made of ~30–40% olivine and 60–70% plagioclase, within horizons of dunites and depleted harzburgites containing 1–10% of interstitial plagioclase crystals. The increase in plagioclase content from the 'impregnated' harzburgites–dunites to the troctolitic segregations is gradational: it occurs over a few centimetres to a few decimetres. The impregnated horizons may reach 10 m in thickness and contain several parallel troctolitic segregations with a characteristic spacing of 1–2 m. The troctolitic segregations themselves generally have a moderate lateral extent (10–30 m at most), but their host impregnated horizons can be followed laterally over a few kilometres. The thick impregnated horizons have a preferred sub-horizontal orientation clearly dis-



FIG 2 a, Field view of three thick horizons of impregnated harzburgites and dunites in the mantle section of the Maqsad area. The horizon in the foreground shows two irregular troctolitic pockets (encircled by dashed lines). Note that the two horizons in the middle distance (encircled by dashed lines) are sub-parallel to the palaeo-Moho (M). These horizons are 1–2 km apart, which corresponds to a spacing of a few hundred metres

along the palaeo-vertical direction. Pale brown outcrops just below the palaeo-Moho represent the 300-m-thick dunitic transition zone between the mantle harzburgites and the layered gabbros. b, Thin section (crossed Nicol prisms) view of a troctolite pocket. Note the recrystallized texture of plagioclase and the pyroxenitic corona around olivine crystals. Field of view, 2 cm.

cordant with respect to the plastic flow structures, and are spaced by a few hundred metres along the palaeo-vertical (Figs 2a and 3).

Olivine in the impregnated horizons has subgrain boundaries indicative of plastic deformation, and highly magnesian compositions (90–92% forsterite) compatible with a residual origin. In the troctolites, olivine generally has an interstitial to poikilitic habit and significantly lower forsterite contents (87–90%), more compatible with precipitation from a melt. Some olivine grains display deformed textural characters and can be interpreted as xenocrysts. Plagioclase generally occurs as mosaics of equant grains 1–3 mm in size with well developed triple junctions (Fig. 2b). Relicts of magmatic plagioclase (elongated subhedral laths up to 1 cm in length) are rarely preserved. Contacts between olivine and plagioclase are underlined by a pyroxenitic corona (Fig. 2b). These features indicate that the troctolites experienced extensive sub-solidus grain-boundary adjustment and intracrystalline diffusion at high temperature. The absence of shear deformation texture in the troctolites indicate purely static recrystallization¹⁹.

Troctolites are also found in dykes with sharp contacts with their host harzburgites. Troctolites within these dykes contain usually minor amounts of clinopyroxene with post-cumulus characteristics (trails of centimetre-sized oikocrysts). Their bulk-rock and mineral chemistry is more evolved than the diffuse troctolites (Fig. 4). Plagioclase recrystallization is less developed in the dykes than in the diffuse troctolites. A continuum exists between diffuse troctolitic segregations and dykes. Troctolites with transitional characters have rather well defined boundaries but display a limited development of impregnated margins. Transitional troctolites are sub-vertical features with no preferred azimuth, whereas troctolite dykes are preferentially oriented parallel to the Maqsad ridge (Fig. 3).

Away from the diapir, olivine gabbro becomes the dominant cumulate lithology. Olivine gabbros exhibit fine- to medium-grained (0.2–2 mm) equigranular adcumulate textures indicative of simultaneous crystallization of plagioclase, olivine and clinopyroxene. Gabbros devoid of olivine are uncommon. The compositions (in terms of both major elements (Fig. 4) and trace elements²³) of olivine gabbros and troctolites are consistent with a common parental liquid of MORB type for these cumulates.

Plagioclase may occur as subhedral laths showing a preferred orientation that can be attributed to alignment during a magmatic flow. Recrystallization is poorly developed, and coronitic textures virtually absent. Olivine gabbros are always found in dykes with

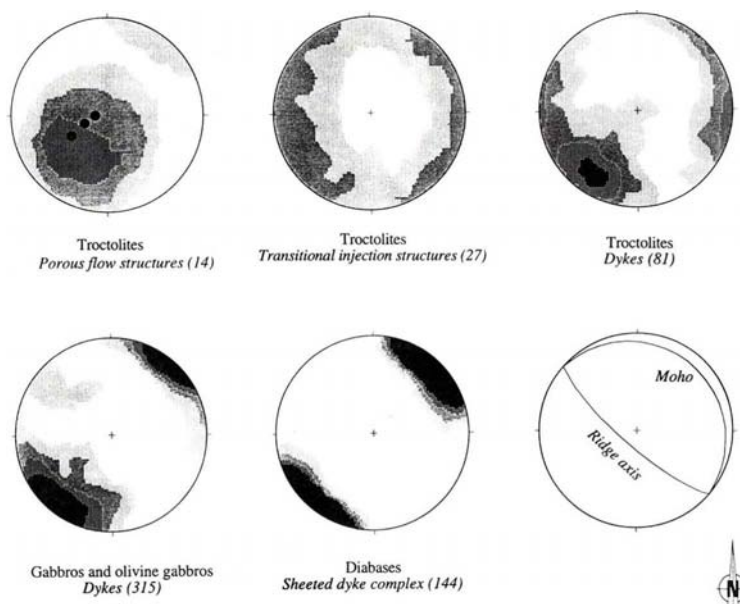
clear-cut boundaries and injection textures diagnostic of hydraulic fracturing¹¹.

In contrast to the troctolite dykes which are generally isolated features, olivine gabbro dykes occur in swarms where they form typically 1–10% of the outcrop surface (Fig. 1). Olivine gabbro dykes are restricted to the periphery of the diapir (that is, the zone of horizontal mantle flow diverging from the upwelling), and show a strong preferred orientation parallel to the ridge axis (Fig. 3). In the diapir itself, they may be found along late fault zones in association with amphibole gabbros, diorites and plagiogranites (Fig. 1).

Gabbro-norite dykes appear in the transition zone between the mantle accreted from the Maqsad diapir and the surrounding lithosphere (Fig. 1). They are characterized by the crystallization of bronzite, Fe–Ti oxides and brown amphibole, indicative of very evolved melt compositions^{24,25}. Textures in these gabbro-norites are either fine grained or pegmatitic (crystals up to several decimetres in size). Fine-grained gabbro-norites occur within centimetre- to decimetre-sized dykes with coarse-grained margins pointing to harisitic growth on the dyke walls. Pegmatitic gabbro-norites occur within swarms of giant dykes (up to 10 m thick). They are found in association with hectometre-sized intrusions of pegmatitic clinopyroxenites and websterites. Pegmatitic gabbro-norites and pyroxenite have been interpreted as the footprint of a massive melt injection event that occurred when the Maqsad diapir impinged on the previously accreted lithosphere, before the opening of the Maqsad ridge^{4,17}.

The present data show that the Maqsad diapir, previously recognized on a structural basis^{1–6}, has a thermal signature imprinted within its cumulate features. Cumulate composition allows the deduction of the melt temperature at the time of crystallization, assuming, in the present case, a dry MORB parental liquid. Cumulate texture can be qualitatively related to the mantle temperature during and after crystallization. The general picture emerging from our observations can be summarized as follows. In the central part of the diapir, the temperature was too high (>1,320 °C; refs 7–9) to allow crystallization of silicate minerals. Dunitic layers, interpreted as former melt channels where orthopyroxene was dissolved^{14,26}, are the only traces of melt migration there. Unfortunately, at such temperatures, the mantle was still actively flowing and these channels were completely deformed and transposed into parallelism with the plastic flow structures. As a consequence, their original structure, orientation, size and spacing have been strongly altered. Moving

FIG. 3 Orientation of cumulate features in the Maqсад area. The orientation of both the palaeo-Moho and the palaeo-ridge axis (deduced from the average azimuth of the diabases in the sheeted-dyke complex) are shown (bottom right) for comparison. Each stereonet is a lower-hemisphere equal-area projection with contours at 1, 2, 3 and >4%. The number of measurements is shown in brackets. Large black dots on the stereonet for porous-flow troctolites (top left) show the orientation of the three thickest structures that we recognized in the Maqсад area. The small number of data for porous-flow troctolites reflects the difficulty of measuring the orientation of these structures rather than their relative abundance. It appears that the orientation of cumulate features presents a well characterized evolution according to their composition and contact relationships with the harzburgites, reflecting the increasing influence of the lithospheric stress field from diffuse troctolites to olivine gabbro dykes.



towards the periphery of the upwelling, the melt temperature decreased slightly but sufficiently to induce fractional crystallization of the melt along the plagioclase-olivine cotectic, resulting in the formation of troctolites (at temperatures of 1,210 to 1,180 °C; refs 7–9). A high mantle temperature, probably equal or very close to the melt temperature, is attested by pervasive static recovery of these cumulates. Plastic flow became, however, very sluggish and the troctolitic cumulates were not sheared after they crystallized. The more evolved troctolites occur within dykes, but the more primitive ones occur as segregations with diffuse boundaries. Porous flow is indicated by the gradational decrease of interstitial cumulus minerals in the peridotite away from these segregations. Porosity cannot, strictly speaking, be quantified from the observation of cumulates. We can at least state that, when the Maqсад diapir solidified, the melt was preferentially distributed in horizons several metres in thickness and spaced by a few hundred metres along the vertical direction. In these horizons, melt had an intergranular distribution but its concentration reached locally very high values, resulting in the formation of decimetre- to metre-sized segregations made of 100% cumulus minerals. This observation is in agreement with some theoretical and experimental results suggesting that melt percolation in the mantle must involve dramatic porosity variations^{27,28}. It is impossible to deduce from field observations if the troctolite segregations and their host impregnated horizons formed in response to vertical melt migration (compaction *sensu stricto*) or to lateral melt injection in porous flow channels, or both, but it is tempting to see there evidence for compaction or decompaction structures whose existence has been suggested by two-phase flow modelling^{29–32}.

At the periphery of the diapir, clinopyroxene joined plagioclase and olivine as a cumulus phase resulting in the formation of olivine gabbros diagnostic of a melt temperature lower than ~1,180 °C (refs 7–9). The field and textural features of olivine gabbros are compatible with melt flow in cracks and crystal settling at a given grain size determined by the Stokes velocity of the cumulus minerals³³. Further away from the diapir, in the surrounding lithosphere, crystallization of evolved and hydrated cumulates point to melt temperatures in the range 1,000–1,100 °C. The textures of these evolved cumulates (for example, pegmatites, and crescumulates on dyke walls) point to rapid crystallization and injection in a mantle significantly cooler than the melt itself³⁴. The lack of textures attributable to melt quenching indicates that Maqсад harzburgites escaped melt injection after they cooled below ~600 °C.

The preservation of the zoned distribution of mafic cumulates around the Maqсад diapir has two independent implications: (1)

the absence of troctolites at the periphery of the diapir implies that high-temperature features were not emplaced at shallow mantle levels (< 2 km below the Moho) in the early stage of the opening of the Maqсад spreading centre, which is consistent with the progressive erosion of an axial lithospheric lid by an asthenospheric diapir; (2) the scarcity of gabbroic cumulates in the diapir itself implies that the melt remaining after troctolite crystallization (~15% crystallization from a MORB) was very efficiently extracted from the diapir before its temperature fell below ~1,180 °C.

We conclude that the brittle-ductile transition for melt trans-

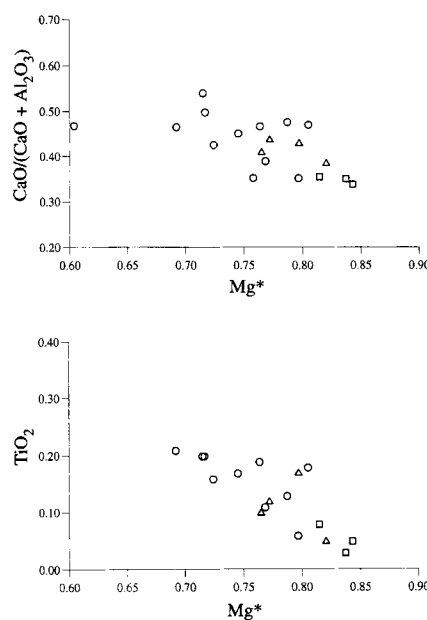


FIG. 4 Bulk-rock major-element variations, showing examples of differentiation trends attributable to fractional crystallization in mafic cumulates of the Maqсад area. Squares, troctolites; triangles, clinopyroxene-troctolites; dots, gabbros. $Mg^* = (MgO)/(MgO + FeO^*)$, where FeO^* is 0.86 times the total FeO . (TiO_2 concentration is given in weight per cents.) These trends reflect the evolution of phase compositions: the forsterite content of olivine ranges from 87–91% in troctolites to 67–87% in olivine gabbros; the anorthite content of plagioclase ranges from 83–93% in troctolites to 68–85% in olivine gabbros; and the ferrosilite content of clinopyroxene ranges from 3.5–4.5% in clinopyroxene-troctolites to 5–9% in olivine gabbros. These mineralogical and geochemical evolutions are typical of MORBs^{20–22}.

port in the mantle occurs in a narrow temperature interval, bracketed, in the case of MORBs, by the crystallization field of troctolites (1,180–1,210 °C; refs 7–9). This temperature is rather high and may explain why dykes dominate the ophiolitic 'landscape', compared to porous flow structures, even if dyking is probably a minor melt-migration mechanism in the partially molten mantle itself. □

Received 7 April; accepted 13 November 1995.

- Rabinowicz, M., Ceuleneer, G. & Nicolas, A. *J. geophys. Res.* **92**, 3475–3486 (1987).
- Ceuleneer, G., Nicolas, A. & Boudier, F. *Tectonophysics* **151**, 1–26 (1988).
- Nicolas, A., Ceuleneer, G., Boudier, F. & Missen, M. *Tectonophysics* **151**, 27–56 (1988).
- Ceuleneer, G. in *Ophiolite Genesis and Evolution of the Oceanic Lithosphere* (eds Peters, T. J., Nicolas, A. & Coleman, R. G.) 149–175 (Kluwer, Dordrecht, 1991).
- Ceuleneer, G. & Rabinowicz, M. in *Mantle Flow and Melt Generation at Mid-ocean Ridges* (eds Phipps Morgan, J., Blackman, D. K. & Sinton, J. M.) 123–154 (Geophysical Monogr. 71, Am. Geophys. Union, Washington DC, 1992).
- Nicolas, A. & Boudier, F. *J. geophys. Res.* **100**, 6179–6198 (1995).
- Green, D. H., Hibberson, W. O. & Jaques, A. L. in *The Earth: Its Origin, Structure and Evolution* (ed. McElhinny, M. W.) 265–290 (Academic, New York, 1979).
- Bender, J. F., Hodeges, F. N. & Bence, A. E. *Earth planet. Sci. Lett.* **41**, 277–302 (1978).
- Bender, J. F., Langmuir, C. H. & Hanson, G. N. *J. Petrology* **25**, 213–254 (1984).
- Dick, H. J. B. *Bull. Ore. St. Dep. Geol. miner. Ind.* **96**, 63–78 (1977).
- Nicolas, A. *J. Petrology* **27**, 999–1022 (1986).
- Ildefonso, B., Nicolas, A. & Boudier, F. *Nature* **366**, 673–675 (1994).
- Kelemen, P. B. *J. Petrology* **31**, 51–98 (1990).
- Kelemen, P. B., Shimizu, N. & Slaters, V. J. M. *Nature* **375**, 747–753 (1995).
- Takahashi, N. *Nature* **359**, 52–55 (1992).
- Amri, I. thesis, Univ. Toulouse (1995).
- Amri, I., Benoit, M. & Ceuleneer, G. *Earth planet. Sci. Lett.* (in the press).
- Leblanc, M. & Ceuleneer, G. *Lithos* **27**, 231–257 (1992).
- Vernon, R. H. Scott, J. *Geology* **6**, 337–351 (1970).
- Dick, H. J. B. et al. *ODP Sci. Res.* **118**, 439–538 (1991).
- Elthon, D., Steward, M. & Ross, D. K. J. *geophys. Res.* **97**, 15189–15200 (1992).
- Batzis, R. & Niu, Y. *J. geophys. Res.* **97**, 6779–6797 (1992).
- Benoit, M., Polvé, M. & Ceuleneer, G. in *Proc. 2nd int. Workshop on Orogenic Lherzolites and Mantle Processes* (eds Barsczus, H. G. & Gervilla, F.) 6 (Instituto Andaluz de Ciencias de la Tierra de Granada, Granada, 1995).
- Juteau, T. & Whitechurch, H. *Bull. Soc. géol. Fr.* **8**, 189–204 (1989).
- Hekini, R., Bideau, D., Hébert, R. & Niu, Y. *J. geophys. Res.* **100**, 10163–10185 (1995).
- Ceuleneer, G. (abstr.) *Eos* **73**, 537 (1992).
- Stevenson, D. *Geophys. Res. Lett.* **16**, 1067–1070 (1989).
- Daines, M. J. & Kohlstedt, D. L. *Phil. Trans. R. Soc. Lond. A* **342**, 43–52 (1993).
- Scott, D. & Stevenson, D. *J. geophys. Res.* **91**, 9283–9296 (1986).
- Spiegelman, M. *Phil. Trans. R. Soc. Lond. A* **342**, 23–41 (1993).
- Sparks, D. W. & Parmentier, E. M. *Earth planet. Sci. Lett.* **105**, 368–377 (1991).
- Khodakovskii, G., Rabinowicz, M., Ceuleneer, G. & Trubitsyn, V. P. *Earth planet. Sci. Lett.* **134**, 267–281 (1995).
- Martin, D. & Nokes, R. J. *Petrology* **30**, 1471–1500 (1989).
- Jahns, R. H. & Burnham, C. W. *Econ. Geol.* **64**, 843–864 (1969).

ACKNOWLEDGEMENTS. This study is based on field work conducted in 1992 and 1993, made possible by the Oman Ministry of Petroleum and Minerals. We thank H. Al Azri for his constant support and interest; P. Kelemen, G. Hirth and M. Spiegelman for reviews; M. Rabinowicz, G. Khodakovskii, M. Benoit, M. Polvé, H. Dick and J.-L. Vigneresse for discussions; M. Bohn for his help during microprobe data acquisition; and A.-M. Roquet for preparing the hundreds of thin rock sections. This work was supported by the Centre National de la Recherche Scientifique under the "Dynamique et Bilan de la Terre" program.

A crustal mineral in a mantle diamond

Leon R. M. Daniels*, John J. Gurney* & Ben Harte†

* Department of Geological Sciences, University of Cape Town, Rondebosch 7700, South Africa

† Department of Geology and Geophysics, University of Edinburgh, King's Buildings, Edinburgh EH9 3JW, UK

RECENT studies of diamonds, inclusions in diamonds, and rock fragments from the Earth's mantle have shown the presence of minerals and rocks with a chemical composition that suggests a previous origin in the crust; presumably, this crustal material was transported into the mantle at a subduction zone. The chemical evidence of previous residence of mantle materials in the crust has come principally from determinations of the proportions of the isotopes of oxygen^{1,2}, sulphur^{3,4}, lead^{4,5} and carbon⁶, which typically show a much wider range of composition in the crust than in material of purely mantle origin. Here we report a less subtle crustal signature in a mantle-derived diamond: an inclusion of the mineral staurolite. Staurolite has

never previously been found in mantle diamonds or rocks, but is a common mineral in metamorphic rocks in the crust. The characteristics of the staurolite suggest not only the probability that it formed in rocks with a crustal bulk composition, but also the possibility that it formed by metamorphism in the crust or in a subduction zone, before being incorporated in the mantle and, eventually, in a growing diamond.

The staurolite inclusion was found in a diamond from the Dokolwayo kimberlite, which is situated in northeastern Swaziland, close to the margin of the Kaapvaal craton, and erupted before the outpourings of the Stormberg volcanics 200 Myr ago⁷, in the Karoo province. Diamond-mining operations have shown that the kimberlitic diatreme consists of at least 11 intrusions. Unfortunately, the near-surface material is extremely weathered and the small number of upper-mantle rocks (xenoliths) that have been found have been severely altered and unsuitable for detailed study. However, the presence of garnet- and chromite-peridotites, eclogites, and megacryst-suite minerals indicating a typical kimberlitic upper-mantle association has been inferred from electron microprobe studies of the chemical composition of large mineral grains (macrocrysts) in the kimberlite⁸. Furthermore, with the exception of the staurolite, the chemical compositions of mineral inclusions within diamonds from Dokolwayo commonly show the same peridotitic and eclogitic assemblages that are found in diamonds from kimberlites worldwide⁹. Thus the Dokolwayo assemblages in general have typical kimberlitic associations indicating an upper-mantle origin, and do not provide evidence of diamond formation within the crust (compare ref. 10). Apart from the unique occurrence of staurolite described here, the only unusual feature of the Dokolwayo diamond assemblages is the relative abundance of coesite⁸; perhaps this, as we shall argue for the staurolite, provides evidence of incorporation of crustal material into the mantle.

The staurolite was recovered from a diamond (specimen D1 3a11) > 3 mm in largest dimension, and the staurolite was completely enclosed by the diamond. No cracks led to the staurolite inclusion, which was present as a discrete crystal of

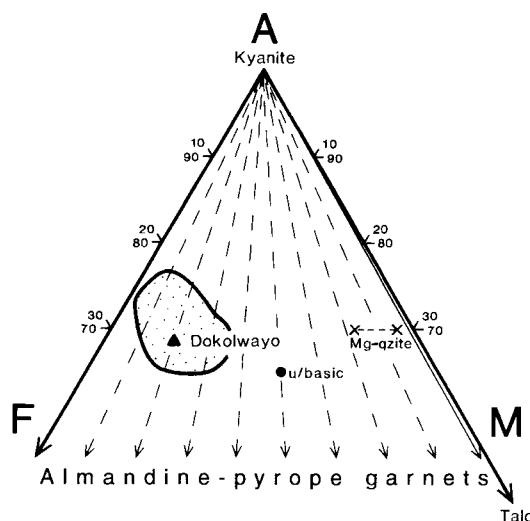


FIG. 1 Compositions of staurolites in AFM (Al_2O_3 –FeO–MgO) molecular projection. The stippled field embraces the compositions of staurolites from pelitic rocks as shown by the compilations of Griffen and Ribbe²¹ and Holdaway *et al.*²³, and also includes staurolite compositions given for eclogite facies pelites^{14,24} and meta-volcanics¹⁸. The composition of the Dokolwayo staurolite is shown within this field. Also shown are the compositions of a staurolite (Table 1, analysis 8) from an Mg-rich ultrabasic (u/basic) and staurolites labelled Mg-qzite (showing Fe–Mg range only) from the Dora Maira magnesian quartzite²⁶. The arrowed tielines from kyanite to garnets (ranging from almandine to 98% pyrope) and talc show the stable assemblages in equilibrium with coesite and phengite in the Dora–Maira rocks²⁴; note that staurolite is not stable in these assemblages, although it may occur in the same rocks as inclusions isolated within garnet (see text).

~80 µm diameter. Thus the staurolite appears to have been included in the diamond at the time of diamond growth. In addition to the staurolite, graphite was recovered as an inclusion from the same diamond. After careful visual examination under a binocular microscope, cleaning the diamond in concentrated HF for 48 hours, and photography, the inclusions were broken out, mounted in epoxy resin on a glass slide and hand-polished. The composition of the staurolite was determined by electron microprobe analysis on a Cameca Camebax instrument, using standards of similar composition to the staurolite and the ZAF correction procedure. The carbon-isotope composition of the diamond was determined by combustion and mass spectrometry, yielding a $\delta^{13}\text{C}$ value of -5.0‰ .

The relatively large size of the diamond containing the staurolite is in the normal range for diamonds of mantle genesis, and contrasts with that for diamonds of crustal origin, which are microdiamonds, typically < 20 µm across occurring as inclusions

in garnet or zircon¹⁰. In addition, the -5.0‰ carbon isotope composition is a typical mantle composition⁶; it does not suggest a crustal origin for the carbon, and is within the range of carbon-isotope compositions obtained for other Dokolwayo diamonds which have characteristic mantle peridotitic and eclogitic inclusions¹¹.

Staurolite normally has a very characteristic and restricted range of occurrences; it most typically occurs in metamorphosed terrigenous sediments of pelitic composition, crystallized under crustal conditions of moderate temperature and moderate-to-high pressure. Thus it is a common mineral in pelitic rocks at medium to high grades in Barrovian and related amphibolite-facies metamorphic sequences^{12,13}, and its range of pressure conditions extends upwards into the eclogite facies¹⁴.

The main departure from the occurrence of staurolite in pelitic protoliths is its presence in metamorphosed basic rocks, mixed pelitic and basic igneous compositions, and hydrothermally altered basic igneous compositions¹⁵⁻¹⁸. The type of hydrothermal

TABLE 1 Staurolite analyses

	Analysis no.								
	1	2	3	4	5	6	7	8	9
Weight per cent									
SiO ₂	27.63	27.62	27.26	27.66	28.4	28.77	27.88	27.42	28.07
Al ₂ O ₃	53.59	54.22	53.62	54.65	53.96	55.46	52.81	54.66	52.32
TiO ₂	0.53	0.61	0.65	0.43	0.32	0.43	0.67	0.23	0.75
Cr ₂ O ₃	0.00	0.04	0.08	n.a.	n.a.	n.a.	n.a.	0.02	1.78
FeO(tot.)*	13.63	13.03	14.26	11.45	10.77	11.28	12.74	9.31	11.19
MgO	2.08	2.24	1.72	1.04	1.52	1.6	2.79	6.48	2.79
MnO	0.34	0.21	0.07	0.28	n.a.	n.a.	0.14	0.03	0.13
ZnO	1.63	0.29	0.25	2.51	0.73	0.75	0.43	0.03	0.9
Li ₂ O	n.a.	0.06	0.13	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
H ₂ O	n.a.	1.78	1.81	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Total	99.43	100.1	99.85	98.02	95.7	98.29	97.46	98.18	97.93
Cations on the basis of 48 oxygens†									
Si	7.698	7.675	7.633	7.753	8.023	7.929	7.834	7.525	7.850
Al	17.597	17.756	17.696	18.053	17.967	18.015	17.489	17.679	17.245
Ti	0.111	0.127	0.137	0.091	0.068	0.089	0.142	0.047	0.158
Cr	0.000	0.009	0.018					0.004	0.394
Fe	3.176	3.028	3.339	2.684	2.545	2.600	2.994	2.137	2.617
Mg	0.864	0.928	0.718	0.434	0.640	0.657	1.169	2.651	1.163
Mn	0.080	0.049	0.017	0.066			0.033	0.007	0.031
Zn	0.335	0.060	0.052	0.519	0.152	0.153	0.089	0.006	0.186
Li		0.067	0.146						
H	3.06‡	3.299	3.381	3.06‡	3.06‡	3.06‡	3.06‡	3.06‡	3.06‡
Total	32.922	32.998	33.137	32.660	32.455	32.504	32.810	33.116	32.703
Total without Li and H	29.862	29.632	29.609	29.600	29.395	29.444	29.750	30.056	29.643
Si + Al	25.296	25.431	25.329	25.805	25.990	25.945	25.323	25.204	25.095
Total R2+	4.455	4.065	4.126	3.704	3.337	3.410	4.285	4.800	3.997
Cations normalized to (Si + Al) = 25.53									
Si	7.770	7.705	7.694	7.670	7.881	7.803	7.898	7.622	7.986
Al	17.760	17.825	17.836	17.860	17.649	17.727	17.632	17.908	17.544
Ti	0.112	0.128	0.138	0.090	0.067	0.088	0.143	0.048	0.160
Cr	0.000	0.009	0.018					0.004	0.400
Fe	3.205	3.040	3.366	2.655	2.500	2.558	3.018	2.164	2.662
Mg	0.872	0.931	0.724	0.430	0.629	0.647	1.178	2.685	1.183
Mn	0.081	0.050	0.017	0.066			0.034	0.007	0.031
Zn	0.338	0.060	0.052	0.514	0.150	0.150	0.090	0.006	0.189
Li		0.067	0.148						
H		3.312	3.408						
Total without Li and H	30.139	29.747	29.844	29.284	28.875	28.973	29.993	30.445	30.157
Si + Al	25.530	25.530	25.530	25.530	25.530	25.530	25.530	25.530	25.530
Total R2+	4.497	4.080	4.158	3.665	3.278	3.355	4.320	4.862	4.066
Estimated H§	2.199			4.052	4.706	4.546	2.301	1.857	1.449

Samples analysed: 1, inclusion in Dokolwayo diamond; 2, 3, relatively low and high Fe/(Fe + Mg) staurolites²³; 4, eclogite facies pelitic staurolite¹⁴; 5, 6, staurolites included in garnets of normal pelite in coesite-bearing unit, Dora Maira²⁴; 7, staurolite in gedrite-staurolite-cordierite-chlorite metavolcanic¹⁸; 8, 9, Mg-rich and Cr-rich (respectively) staurolites from ultrabasic and basic protoliths²⁰. n.a., not analysed.

* Total Fe as FeO.

† Assuming 3.06 H cations if no H₂O wt% available.

‡ Assumed number of H cations (see ref. 23).

§ Calculated as deficit of cation charges using cations analysed for and assuming a total of 96 positive charges.

|| Total divalent cations.

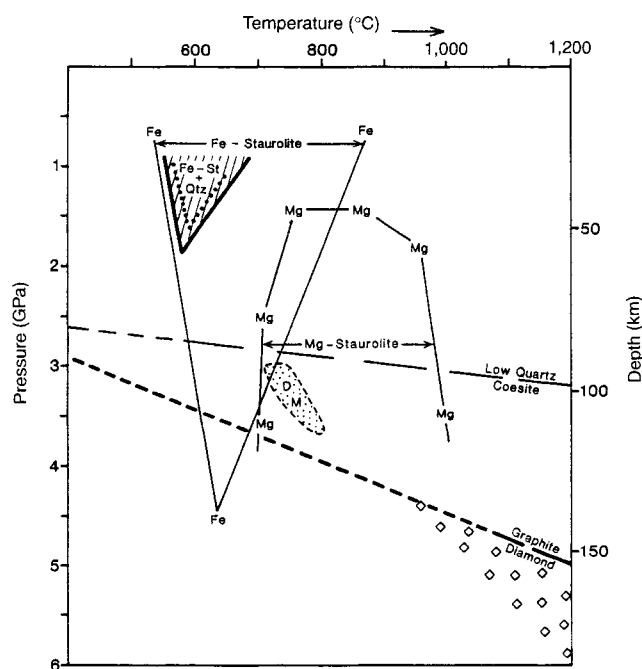


FIG. 2 Stability fields for diamond, coesite and various staurolite-bearing assemblages. The diamond-graphite transition boundary, extrapolated to low temperatures, is from ref. 35, and that for coesite from reference 36. The absolute stability limits (not involving the presence or absence of quartz) of respectively Fe-staurolite and Mg-staurolite are shown with Fe and Mg labels and are taken from refs 30 and 34. The lined field shows that the upper pressure stability field for Fe-staurolite and quartz, as given by Ganguly³⁰; Ganguly's field provides a maximum pressure estimate of this field because the experimental data of Rao and Johannes³¹ place the high-pressure limits of this assemblage ~ 0.3 GPa lower (very close to the lines of heavy dots). The heavy dots actually show the high-pressure limit of staurolite + quartz in normal pelitic compositions (with high Fe/(Fe + Mg) staurolite) as estimated for eclogitic assemblages¹⁴. The dotted field labelled DM gives the P - T conditions estimated for the formation of assemblages in equilibrium with coesite in the coesite-bearing unit of Dora-Maira^{24,27,28}. The field with diamond symbols indicates the P - T conditions of formation in the diamond stability field of mantle peridotite rocks associated with kimberlitic diamonds in the Kaapvaal craton^{37,38}.

alteration concerned is commonly associated with igneous activity in oceanic crustal settings, and leads to the formation of rocks enriched in Fe, Mg and Al¹⁵⁻¹⁸. Upon subsequent metamorphism such rocks often form mineral assemblages whose most widespread characteristic is the occurrence of Fe-Mg amphiboles, but they also have assemblages containing minerals like staurolite and hornblende¹⁵⁻¹⁸. These mineral assemblages often contain quartz, but this is not a universal characteristic as in the case of straightforward pelitic assemblages. Much more rarely, staurolite occurs in crustal metamorphic basic and ultrabasic igneous rocks not previously extensively modified by pervasive alteration^{19,20}. In these cases, metamorphic reconstitution is often incomplete and the staurolite is present in small amounts in association with hornblende and usually without quartz.

The chemical composition of the staurolite from Dokolwayo is given in Table 1, together with representative analyses of staurolites from crustal metamorphic rocks. Figure 1 further illustrates general compositional features. Unfortunately, we have been unable to analyse the Dokolwayo staurolite for H and Li, which are respectively essential and sometimes important elements within staurolite (see refs 21-23 and references therein), and which will affect the stability relationships with respect to other phases. A plan to analyse these elements by ion microprobe has been thwarted by the misplacement of the staurolite inclusion. To enable more detailed comparison of the staurolite chemical composition we have calculated cation contents for the staurolites

in Table 1. Following the recommendations of Holdaway *et al.*^{22,23}, we have calculated cation contents in two ways: (1) by assuming a given amount of H (3.06 cations of H per 48 oxygen formula unit); and (2) by assuming a combined total of 25.53 Si + Al cations. Both methods yield similar results for the cations measured and a moderately high cation total for the Dokolwayo staurolite (Table 1), which suggests that there is not an unusually high content of unanalysed species.

The Dokolwayo staurolite shows a Fe-rich composition with a striking resemblance to that of staurolites from normal pelitic rocks. For comparison, Table 1 (analyses 2-6) gives pelitic staurolite compositions of relatively high Fe/(Fe + Mg) and low Fe/(Fe + Mg) from the compilation of Holdaway *et al.*²³, together with an eclogite facies pelitic staurolite of median Fe/(Fe + Mg) (ref. 14), and the composition of two staurolite inclusions within garnets in normal (Fe-rich) pelites from the coesite-bearing unit eclogitic rocks of Dora-Maira²⁴. The tabulation shows that the Dokolwayo staurolite lies within the range of these staurolite compositions. Other metamorphosed bulk compositions such as those of basic igneous rocks, and particularly those of chemically altered basic rocks and mixed pelitic-basic compositions, often yield staurolites (for example, Table 1, analysis 7) very similar in composition to those from pelites. It is notable that the total range of staurolite compositions in pelites and altered metabasites is very restricted if one ignores occasional specimens showing very high concentrations of trace/minor elements such as Zn, Li and Co^{21,23}. Figure 1 shows that the Dokolwayo staurolite plots well within this restricted composition field.

However, staurolites from metamorphosed basic and ultra-basic protoliths may show a trend towards obviously Mg- and Cr-rich compositions²⁰, as shown by analyses 8 and 9 in Table 1, where the high-Mg specimen is from a meta-troctolite and the high-Cr staurolite occurs in a meta-tuff. The occurrence of such compositions suggests that high Mg and Cr characteristics might be expected if staurolite formed in typical mantle rock compositions. In addition the possibility for forming high-Mg staurolite in more Mg-rich bulk rock compositions at high pressure is shown by some experimentally synthesized staurolites²⁵ and by some other natural occurrences³⁹, especially the magnesian staurolite inclusions (plotted in Fig. 1) in pyrope garnet from the magnesian quartzite of the Dora-Maira terrain^{24,26}. The Dokolwayo diamond inclusion is clearly not a member of such Mg-rich staurolite parageneses.

Thus by comparison with other natural occurrences, the compositional characteristics of the Dokolwayo staurolite inclusion strongly suggest formation in a pelitic, altered basic or mixed pelitic-basic composition, by metamorphism under moderate-to-high pressures and moderate temperatures in the crust. It is notable that there is no evidence for formation of such Fe-rich staurolite at exceptionally high crustal pressures. This is illustrated by the mineral assemblages found in the ultra-high-pressure, coesite-bearing, Dora-Maira terrain. Here the critical assemblage found in equilibrium with free SiO₂ (coesite) in appropriate pelitic compositions is garnet + kyanite (together with phengite and jadeite, as well as coesite)^{24,27}. As illustrated, in Fig. 1, the garnet + kyanite + coesite assemblage covers virtually the whole range of Fe/Mg ratios in SiO₂ saturated bulk compositions, except in the most extreme magnesian assemblages where talc also occurs. Thus there is no stability field for staurolite in any bulk composition carrying the SiO₂ phase. Although Fe-rich staurolite (Table 1, analyses 5 and 6) does occur in these same rocks, it does so only as isolated inclusions in garnet, which are not in equilibrium with the coesite; the staurolite having formed at an earlier period during the prograde evolution of the rocks, and before the attainment of coesite stability conditions²⁴. The conditions of formation of the Dora-Maira coesite-bearing assemblages are well constrained^{24,27,28} at 3.0-3.7 GPa at 700-800 °C, and are at shallower depths than the stability field of diamond (Fig. 2).

Experimental data on the stability limits of staurolite and diamond, extrapolated across the pressure-temperature (P - T) conditions where both minerals might stably co-exist, are sum-

marized in Fig. 2. Considerable work has been done on determining the P - T stability limits of typical pelitic staurolite-bearing assemblages, both by experimental studies²⁹⁻³² on Fe-staurolite (that is, pure Fe-staurolite lacking any Mg), and by consideration of the limiting reactions of natural assemblages with Fe-rich staurolites^{12-14,33} (with compositions like those of analyses 1-7, Table 1). In these studies the focus has been upon staurolite coexisting with quartz (SiO_2), and it may be seen from the shaded field in Fig. 2 that the upper pressure stability limit for such staurolite is ~ 1.9 GPa, well away from the stability field for diamond. It is significant that in natural assemblages formed near this upper pressure limit for Fe-staurolite + quartz, Fe-Mg partitioning between staurolite and garnet favours the staurolite as the more Fe-rich phase¹⁴, and thus the substitution of some Mg for Fe in staurolite and garnet will not extend the stability limits of staurolite beyond those of pure Fe-staurolite and quartz¹⁴ (as illustrated by the dotted stability limit in Fig. 2).

If assemblages lacking free SiO_2 are considered, then the total stability field of Fe-staurolite becomes pertinent. The extent of this field, based on Ganguly³⁰, is shown in Fig. 2, but it has to be noted that the P - T location of this field is much more uncertain than that of Fe-staurolite + quartz. As illustrated, the Fe-staurolite stability field overlaps with the extrapolated diamond stability field in a small region of P - T space. When Mg-staurolite is considered then the experimental data³⁴ indicate an extensive field of stability to very high pressures (Fig. 2), and in the absence of a SiO_2 phase, Fe-Mg staurolite solid-solution compositions may have a wide stability field overlapping with the stability fields of coesite and diamond. This wide stability field for staurolite in the absence of a SiO_2 phase (quartz or coesite), would aid the survival of a staurolite crystal formed at lower pressures if it was subsequently subjected to higher pressure conditions. Thus the early-formed staurolite in the Dora-Maira coesite-bearing rocks, noted previously, probably survived alteration because it became isolated from the coesite by becoming enclosed in garnet. If the Dokolwayo staurolite had initially formed in a crustal rock composition containing quartz or coesite, similar entrapment in another phase may have aided its survival as it was carried to higher-pressure conditions and eventually passed into the diamond stability field. Alternatively, the Dokolwayo staurolite may have escaped the possibility of breakdown reactions with a SiO_2 phase by forming in a more basic or altered basic crustal rock composition which lacked an SiO_2 phase in its matrix.

Despite the potentially wide stability field for Fe-Mg staurolite solid solutions, Fig. 2 also emphasizes the unlikelihood of forming an Fe-rich staurolite at the P - T conditions of mantle rocks found together with diamonds of kimberlitic affinity in the Kaapvaal craton and elsewhere. The estimated P - T conditions of formation of such mantle rocks in the diamond stability field (diamond symbols in Fig. 2) overlap with only the highest temperature and pressure conditions of Fig. 2, where only Mg-rich staurolite could be expected to form.

Thus the size, composition and associations of the Dokolwayo diamonds all indicate growth in mantle conditions commonly associated with kimberlitic diamonds, rather than in crustal conditions. On the other hand, the compositional characteristics of the included staurolite suggest formation within a crustal protolith, which is likely to have had the bulk composition of a pelite, a mixed pelitic-basic rock, or a chemically altered basic rock from the ocean floor; the staurolite probably formed during metamorphism of this protolith under crustal conditions at pressures below the stability field of diamond.

The occurrence of the Dokolwayo staurolite inclusion in diamond therefore shows the incorporation of material of crustal provenance within the mantle, and adds to the growing body of empirical evidence for such material, which has come mainly from isotope studies¹⁻⁶. Added significance attaches to the Dokolwayo staurolite inclusion, because it suggests the presence within the mantle of an actual crystal which had grown within the crust or in crustal material in a subduction zone, before being incorporated

in the mantle and then preserved within a diamond which grew under typical mantle conditions. □

Received 13 September; accepted 13 November 1995.

- Jagoutz, E., Dawson, J. B., Hoernes, S., Spetall, B. & Wänke, H. *Lunar planet. Sci.* **15**, 395-396 (1984).
- Ongley, J. S., Basu, A. R. & Kyser, T. K. *Earth planet. Sci. Lett.* **83**, 80-84 (1987).
- Chaussidon, M., Albarède, F. & Sheppard, S. M. F. *Nature* **330**, 242-244 (1987).
- Eldridge, C. S., Compston, W., Williams, I. S., Harris, J. W. & Bristow, J. W. *Nature* **353**, 649-653 (1991).
- Rudnick, R. L., Eldridge, C. S. & Bulanova, G. P. *Geology* **21**, 13-16 (1993).
- Kirkley, M. B., Gurney, J. J., Otter, M. L., Hill, S. J. & Daniels, L. R. *Appl. Geochem.* **6**, 477-474 (1991).
- Allsopp, H. L. & Roddick, J. C. in *Petrogenesis of the Volcanic Rocks of the Karoo Province* (ed. Erlank, A. J.) 267-271 (Spec. Publ. 143, Geol. Soc. South Afr., 1984).
- Daniels, L. R. M. & Gurney, J. J. in *Kimberlites and Related Rocks* Vol. 2 (eds Ross, J. et al.) 1012-1021 (*Proc. 4th Int. Kimberlite Conf.*, Blackwell, London, 1989).
- Meyer, H. O. A. in *Mantle Xenoliths* (ed. Nixon, P. J.) 501-522 (Wiley, Chichester, 1987).
- Sobolev, N. V. & Shatsky, V. S. *Nature* **343**, 742-746 (1990).
- Daniels, L. R. M. thesis, Univ. Cape Town (1991).
- Albee, A. L. *Am. J. Sci.* **263**, 512-536 (1965).
- Harte, B. & Hudson, N. F. C. in *The Caledonides of the British Isles - Reviewed* (eds Harris, A. L., Holland, C. H. & Leake, B. E.) 323-337 (Spec. Publ. 8, Geol. Soc. Lond., 1979).
- Baileve, M., Pinardon, J. L., Kienast, J. R. & Vuichard, J. P. *J. Petrology* **30**, 1321-1349 (1989).
- Robinson, P. & Jaffe, H. W. *Am. J. Sci.* **267**, 389-421 (1969).
- Hudson, N. F. C. & Harte, B. *Am. J. Sci.* **285**, 224-266 (1985).
- Schumacher, J. C. *Am. J. Sci.* **288**, 619-663 (1988).
- Spear, F. S. *J. Petrology* **23**, 383-426 (1982).
- Gibson, G. M. *Mineralog. Mag.* **42**, 153-154 (1978).
- Ward, C. M. *Am. Miner.* **69**, 532-540 (1984).
- Griffin, D. T. & Ribbe, P. H. *Am. J. Sci.* **273A**, 479-495 (1973).
- Holdaway, M. J. *Am. Miner.* **71**, 1135-1141 (1986).
- Holdaway, M. J., Dutrow, B. L. & Shore, P. A. *Miner.* **71**, 1142-1159 (1986).
- Chopin, C., Henry, C. & Michard, A. *Eur. J. Mineral.* **3**, 263-291 (1991).
- Hellman, P. L. & Green, T. H. *Contr. Miner. Petrol.* **68**, 369-377 (1979).
- Chopin, C. *Phil. Trans. R. Soc. Lond.* **321**, 183-197 (1987).
- Scherl, H.-P., Schreyer, W. & Chopin, C. *Contr. Miner. Petrol.* **108**, 1-21 (1991).
- Sharp, Z. D., Essene, E. J. & Hunziker, J. C. *Contr. Miner. Petrol.* **114**, 1-12 (1993).
- Richardson, S. W. *J. Petrology* **9**, 467-489 (1968).
- Ganguly, J. *J. Petrology* **13**, 335-365 (1972).
- Rao, B. B. & Johannes, W. *Nelles Jb. Miner. Mh.* **H10**, 437-444 (1979).
- Dutrow, B. L. & Holdaway, M. J. *J. Petrology* **30**, 229-248 (1989).
- Koons, P. O. & Thompson, A. B. *Chem. Geol.* **50**, 3-30 (1985).
- Schreyer, W. *Mineralog. Mag.* **36A**, 1-26 (1988).
- Kennedy, C. S. & Kennedy, C. G. *J. geophys. Res.* **81**, 2467-2470 (1976).
- Bose, K. & Ganguly, J. *Am. Miner.* **80**, 231-238 (1995).
- Carswell, D. A. & Gibb, F. G. F. *Contr. Miner. Petrol.* **97**, 473-487 (1987).
- McKenzie, D. *Earth planet. Sci. Lett.* **95**, 53-72 (1989).
- Chopin, C. & Sobolev, N. V. in *Ultrahigh Pressure Metamorphism* (eds Coleman, R. G. & Wang, X.) 96-131 (Cambridge Univ. Press, 1995).

ACKNOWLEDGEMENTS. We thank M. Holdaway and O. Navon for reviews, and C. Chopin for helpful comments.

A *Dryopithecus* skeleton and the origins of great-ape locomotion

Salvador Moyà-Solà & Meike Köhler

Institut de Paleontologia Miquel Crusafont, c/Industrial, 23, 08201 Sabadell, Spain

THE evolution of skeletal adaptations to orthograde postures, characteristic of extant hominoids, is of great interest as it provides the key to understanding the origins of apes and humans. We report here the recent discovery of an extraordinary partial skeleton of *Dryopithecus laietanus* from Can Llobateres (Spain). It provides evidence that orthograde postures and locomotion appeared at least 9.5 million years ago¹. Our results indicate that the body structure of this Miocene ape closely resembles that of extant hominoids^{2,3} and differs from the pronograde pattern of Miocene proconsulids^{4,5} in a set of important morphological characters. *Dryopithecus* also shows more traits reflecting structural adaptations for suspension than occurs in African apes. A similar positional behaviour is inferred for *Sivapithecus indicus*, thus strengthening previous hypotheses linking both Miocene forms with *Pongo*⁶⁻⁹.

Associated skeletal material of Miocene apes is scarce. Between *Proconsul*¹⁰ (18 million years (Myr)) and "Lucy"

(3.1 Myr) only the enigmatic specimen of *Oreopithecus*^{11–13} is known. In the 1992–94 field seasons at the Late Miocene site Can Llobateres (near Sabadell, Spain) we recovered numerous postcranial elements of the European ape *Dryopithecus laietanus* (Figs 1 and 2). All bones were found in the same horizon from which the partial cranium CLI 18000 had been recovered previously^{8,9}. The taphonomic context suggests that all material, including the cranial specimen, belongs to a single adult male, dismembered by carnivores and dispersed over an area of at least 900 m². There is no duplication of elements, some of the pieces were found next to articulated bones, and (with the exception of turtle and rhino) there is no trace of other vertebrates in this level. Palaeomagnetism indicates an age of 9.5 Myr¹.

This exceptional find of a partial skeleton of a late Miocene ape provides evidence concerning the origins of modern ape skeletal anatomy and locomotion. Extant anthropoids show two basic patterns of skeletal structures related to locomotion. The first is exemplified by the generalized platyrrhine/cercopithecoid skele-

ton, adapted for pronograde postures, with a long and flexible spinal column, a narrow thorax with laterally situated scapulae and a moderate intermembral index. The second is the modern hominoid skeleton, adapted for orthograde postures, with a shorter and stiffer spinal column, a broader thorax with more dorsally shifted scapulae and a relatively high intermembral index^{4,5}.

Besides *Oreopithecus*, no fossil evidence has been found documenting hominoid skeletal anatomy and locomotion of modern aspect earlier than the oldest *Australopithecus*. The most recent review⁵ of Miocene hominoid postcranial anatomy concluded that Middle to Late Miocene apes, with the exception of *Oreopithecus*, were generalized quadrupeds, structurally similar to *Proconsul*, perhaps with some climbing and suspensory abilities. In contrast to these interpretations, examination of the new skeletal elements reported here indicates that *Dryopithecus* differed markedly from *Proconsul* and the remaining Miocene forms that possessed a monkey-like pattern.

Several features of the trunk of *Dryopithecus* suggest adaptation for orthograde postures. The lumbar vertebrae are proportionally shorter than those of cercopithecoids and proconsulids^{4,14}. The transverse processes originate on the pedicle dorsolaterally on the vertebral body, and the caudally directed spinal processes indicate reduced mobility of the lumbar region. The more dorsally situated costal foveae of the thoracic vertebra imply a more ventral

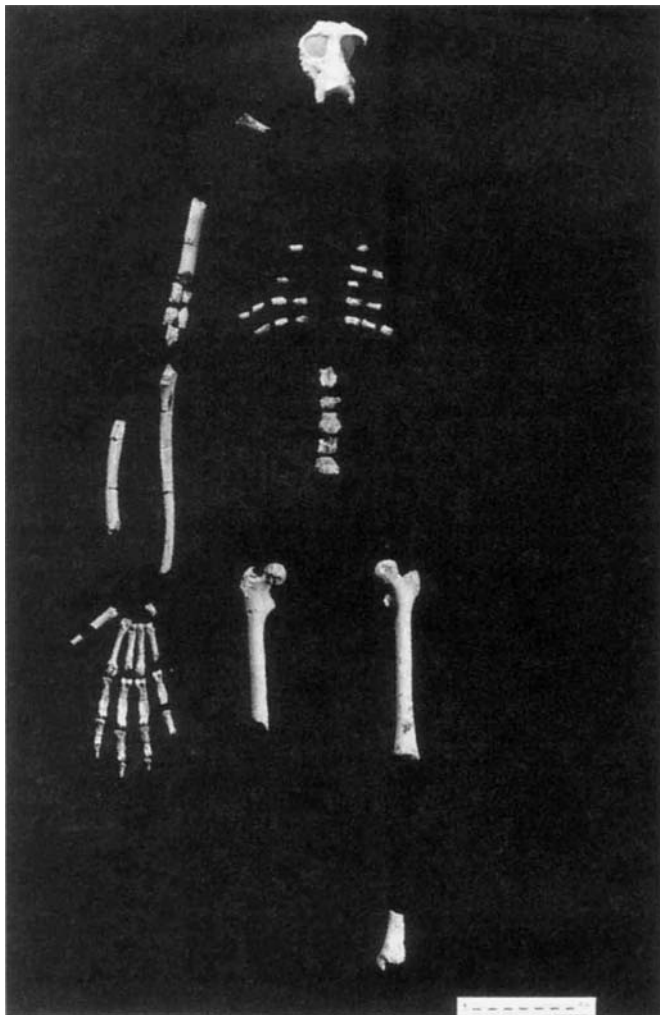


FIG. 1 The *Dryopithecus* skeleton CLI-18800 from Can Llobateres. Forelimb: well-preserved humeral diaphysis (~105 mm long) with insertion areas of deltoideus, pectoralis, teres major and coracobrachialis, other small humeral fragments; ulna nearly complete from the middle of the trochlear notch to the distal end of pronator quadratus insertion; radial diaphysis; hand: triquetrum, all metacarpals, all phalanges except apical ones of the first and fifth fingers, four sesamoids, all of them of the right side. Shoulder girdle: fragment of right clavicle (~40 mm long), showing the sigmoidal curvature from the middle of the shaft to conoid tubercle; trunk: several rib fragments, three incomplete lumbar vertebrae, one thoracic vertebra, two neural arches, and other small fragments; hindlimb: two femora, the left one from caput to the distal end of diaphysis, the right one from caput to the distal third of the shaft; roughly the distal third of tibia.



FIG. 2 The hand of CLI-18800 (palmar view). The hand is very large compared with body size. Relative to the length of the forelimb, its size is comparable to that of the hands of extant large hominoids, particularly *Pan* and *Pongo*. It exhibits a mixture of both primitive hominoid and modern hominoid features. Primitive hominoid characters are the short metacarpals, small articular surfaces of the metacarpals as well as of proximal and medial phalanges and the proximal articular facet of metacarpal 5 with dorsal expansion for the hamate (but flat as in hominoids). These features suggest palmigrady with the transversal axis of the hand held parallel to the support. The modern hominoid characters are the longer second metacarpal and shorter fourth one, and the stabilizing carpo-metacarpal articulation. Further, there are some specialized features typical of suspensory forms, such as long phalanges with strong and distally placed flexor insertions. The robust first metacarpal indicates that the hand was adapted for powerful grasping.

TABLE 1 Bone length of *Dryopithecus laietanus*

	<i>P. nyanzae</i>	<i>D. laietanus</i>	<i>O. bamboli</i>	Orang-utan	Bonobo	Chimp	Baboon
BW (kg)	34	34	32	32.7	31	31.5	30.9
Humerus L		297	298	327–329	282	251–271	244
Radius L		280	283	323	261	235–242	262
Femur L	295	274	255	244–251	285	252–273	281
Tibia L		236(1)	235	215–219	238	214–227	259
IMI		114	118.5	140	104	103	94
H/F		108	117	132	99	99	87
BRAI		94	95	99	92	93–89	107

Bone length of *Dryopithecus laietanus* from Can Llobateres (CLI-18800) placed in a narrow allometric context of catarrhines of similar body mass in order to compare body proportions and probable locomotor adaptations^{20,21}. IMI, Intermembral index; H/F, humero-femoral index; BI, brachial index. Calculation of bone length of *Dryopithecus*: The length of the clavicle is estimated by its midshaft circumference, based on the correlation in catarrhines between clavicle length and midshaft circumference ($r^2 = 0.862$). The average obtained is 161 mm, with a minimum of 138 mm and a maximum of 183 mm. The minimum length of the *Dryopithecus* clavicle, corrected by body mass ($L/\text{cube root of estimated body weight} (=34,000 \text{ g})$) gives a value of 4.2, even slightly superior to that of *Pongo* females (3.9) and clearly superior to that of *Pan* and *Gorilla* (2.8). The length of the humerus (297 mm) is calculated from the reconstructed length of the humerus from St Gaudans (276.2 mm) according to the difference in the circumferences of the two humeri at the level of the distal point of the deltoid impression. The length of the radius (280 mm) is calculated from the reconstructed length of the ulna (from the middle of the trochlear notch to the distal articular surface). The length of the tibia, used only for the estimation of the intermembral index, is calculated as a function of femoral length following the data of Aiello¹⁸. The body weight of *Dryopithecus* was calculated from various postcranial joint surfaces: femur head diameter²⁸ (34 kg), femur head surface²⁹ (37 kg), tibiotalar joint surface³⁰ (34 kg) and articular surface of lumbar vertebral body¹⁴ (30 kg). These dimensions suggest a weight of 34 kg for this individual. The length of the femur and the body weight of *Proconsul nyanzae* is taken from ref. 28. Data for *Oreopithecus* as well as for the remaining catarrhines are taken from ref. 13.

position of the spinal column and hence a broad thorax². The clavicle is proportionally longer than in African apes and comparable in relative size to the lengths of clavicles of *Pongo* and *Hylobates* (Table 1). This suggests that the scapula of *Dryopithecus* was situated dorsally on the thorax as in extant hominoids, and that the clavicle would have been oriented more vertically than in African apes, and similar to *Pongo* and *Hylobates*, a character linked to suspensory postures^{2,3}.

Several characters indicate that the arms were powerful, with a wide range of movements (abduction, flexion, pronation), evoking a picture of habitual climbing and suspension. The humeral shaft is straight, with a slightly convex deltoid plane. The radius is strongly curved and the ulna preserves the insertion area for a powerful brachialis. The triquetrum has a convex surface for the ulnar styloid process, suggesting a reduced stylotriquetral contact³. The hand is very large compared with body size (Fig. 2). The phalanges are long, with strong and distally placed flexor insertions. The robust first metacarpal indicates that the hand was adapted for powerful grasping. The large femoral head relative to neck size and the high femoral neck-shaft angle (133°) reflect a wide range of hip mobility, including abduction¹⁵. Distally, the femoral shaft is notably flattened anteroposteriorly, suggesting abduction as a common position.

Analysis of percentage prediction errors (Fig. 3) and narrow allometry (comparison with other catarrhines of comparable body weight Table 1) of the lengths of long bones^{16,17} indicate that the humerus and radius of *Dryopithecus* are longer and the femur shorter than expected for an African ape of comparable size. As in *Oreopithecus*¹³, *Dryopithecus* exhibits a tendency towards arm elongation and leg reduction, although to a limited degree. Among great apes, this tendency reaches its maximal expression in *Pongo*. Taking into account the existence of a size-related trend towards increasing intermembral index (related to maintenance of competence in climbing^{18–21}) and the fact that African apes possess the proportions expected for their body mass^{3,22}, the proportions observed in *Dryopithecus* can be interpreted as reflecting adaptation to climbing and suspension, which thus seem to have been more dominant in *Dryopithecus* than in modern African apes.

There is only one known Miocene hominoid, besides *Dryopithecus*, with a body structure adapted for orthograde postures: *Oreopithecus*, an endemic insular form from Italy^{11–13}. The morphology of the lumbar vertebrae, long bones and intermembral proportions is similar in the two forms, suggesting comparable locomotor behaviour. Nevertheless, the peculiar cranial and dental anatomy of *Oreopithecus*^{11–13} renders it difficult to appreciate the phylogenetic implication of the resemblances in

postcranial anatomy.

The hypothesis that *Dryopithecus* was a member of the clade of extant great apes was recently strengthened by cranial evidence^{8,9,23,24}, but its position within either the Asian^{8,9} or the African^{23,24} subclade is still a matter for debate. The postcranial skeleton possesses a set of derived characters apparently shared with *Pongo* and absent in African apes (very long arms and short legs relative to its body size, large clavicle, long and curved proximal phalanges with distally placed flexor insertions, distal shaft of femur anteroposteriorly compressed), thus reaffirming the position of this genus in the *Pongo* clade. Nevertheless, *Dryopithecus* shows several primitive hominoid features, namely, short metacarpals, small articular surfaces of metacarpals and phalanges and somewhat elongated lumbar vertebrae. This suggests that the postcranial skeleton of *Dryopithecus* as well as its locomotor behaviour might have been less specialized than in *Pongo* (possibly including a higher degree of quadrupedalism than in the latter⁵).

Having come thus far, it is necessary to examine whether or not the genus *Sivapithecus* belongs to this clade. Recent studies question^{5,25} the hypothesis that linked *Sivapithecus* and *Pongo*^{6,7}. This change of mind came with the report on two almost complete humeri attributed to *Sivapithecus parvada* and *Sivapithecus indicus*, both described as quite robust, retroflexed, medially inclined, with a flat deltoid plane and without marked humeral torsion²⁵, morphology that was taken to indicate that this genus was quadrupedal, similar to *Proconsul*, with a vertically placed scapula, a mediolaterally compressed thorax, a monkey-like proximal and ape-like distal humerus⁵ and a set of cranial features convergent with *Pongo*. The morphology of the humerus of *Sivapithecus parvada* certainly fits this description. In our opinion (based on our reconstruction), the crushed and somewhat distorted²⁵ smaller humerus of *Sivapithecus indicus* differs notably from the latter but closely resembles that of *Dryopithecus fontani*²⁶. It is slender, not retroflexed, its lateral side is almost straight and there is no indication that the deltoid plane was flat. Therefore, it does not contradict the proposed close relationship to *Pongo*, suggested by a large series of apparently shared cranial characters^{6,7}. The genus *Sivapithecus* would thus probably include two different locomotor adaptations: mainly climbing and suspension²⁷ (*Sivapithecus indicus*) and a secondarily more quadrupedal locomotion (*Sivapithecus parvada*)^{5,25}.

Both *Dryopithecus* and *Sivapithecus* possess several primitive hominoid features for which modern great apes share an apparently derived condition. If *Dryopithecus* and *Sivapithecus* are indeed members of the *Pongo* clade, the common ancestor of all

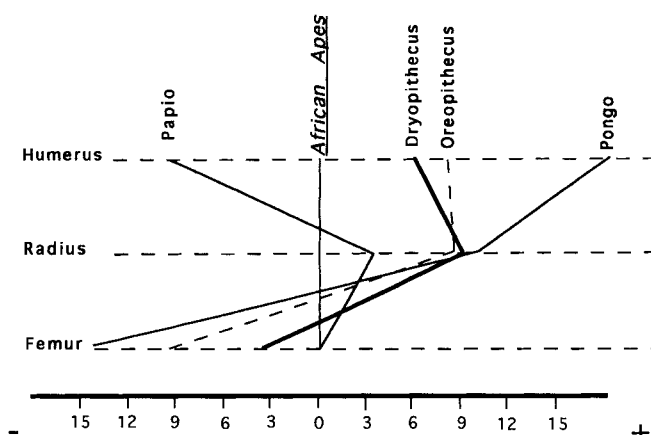


FIG. 3 Percentage prediction errors for the lengths of humerus, radius and femur of *Dryopithecus laietanus* from Can Llobateres with respect to African apes baseline. The deviations (residuals of each species from least-squares regression lines of the scaling between long bone length and body weight) are examined for the lengths of humerus, radius and femur. The percentage deviations of observed values are calculated with respect to the African apes baseline (from equations calculated by Jungers¹⁹). Positive and negative percentage deviations are computed as follows¹⁹: observed value – predicted value/predicted value $\times 100$. This figure expresses graphically positive (longer bone) and negative (shorter bone) deviations. *Dryopithecus* shows not only the key character of orthograde hominoids (intermembral index above 100), but also a degree of forelimb elongation and hindlimb reduction superior to that of African apes and close to that of *Pongo*.

extant great apes would have been more primitive than hitherto inferred by the analysis of extant forms. If so, then resemblances between the Asian and the African subclades in such derived features must reflect homoplasy.

We suggest that Asian and African great apes underwent divergence in their locomotor strategies following their separation from a common ancestor, presumed to be a generalized (orthograde) climber. Asian great apes became adapted to suspension and slow climbing, already present in *Dryopithecus*, whereas the African clade specialized in terrestrial quadrupedalism (*Pan*, *Gorilla*) and bipedalism (*Australopithecus*, *Homo*). The two alternative strategies appeared in independent geographical areas, reflecting the role of geographical isolation as a factor underlying the main dichotomy in the evolution of great apes. □

Received 7 June; accepted 20 October 1995.

- Agusti, J. et al. *J. hum. Evol.* (in the press).
- Enikson, G. E. *Symp. zool. Soc. Lond.* **10**, 136–164 (1963).
- Andrews, P. & Groves, C. in *Gibbon and Siamang* 167–218 (Karger, Basel, 1976).
- Ward, C. V. *Am. J. phys. Anthropol.* **92**, 291–328 (1993).
- Rose, M. D. *J. hum. Evol.* **26**, 387–412 (1994).
- Andrews, P. & Cronin, J. E. *Nature* **297**, 541–546 (1982).
- Pilbeam, D. *Nature* **295**, 232–234 (1982).
- Moya Solà, S. & Köhler, M. *Nature* **365**, 543–545 (1993).
- Moya Solà, S. & Köhler, M. *J. hum. Evol.* **28**, 101–139 (1995).
- Napier, J. R. & Davis, P. R. *The Fore-limb Skeleton and Associated Remains of Proconsul africanus* 1–69 (British Museum Natural History, London, 1959).
- Hürzeler, J. *Ann. Paleont.* **54**, 195–233 (1968).
- Harrison, T. J. *J. hum. Evol.* **15**, 541–583 (1986).
- Jungers, W. L. *J. hum. Evol.* **16**, 445–456 (1987).
- Sanders, W. J. & Bodenbender, B. E. *J. hum. Evol.* **26**, 203–238 (1994).
- Ward, C. V., Walker, A., Teaford, M. F. & Odiambro, I. *Am. J. phys. Anthropol.* **90**, 77–111 (1993).
- Jungers, W. L. *J. hum. Evol.* **16**, 445–456 (1987).
- Jungers, W. L. in *Body Size in Mammalian Paleobiology. Estimation and Biological Implications* (eds Damuth, J. & MacFadden, B. J.) 107–118 (Cambridge Univ. Press, 1990).
- Aiello, L. *Symp. zool. Soc. Lond.* **48**, 331–358 (1981).
- Jungers, W. L. in *Size and Scaling in Primate Biology* (ed. Jungers, W.) 345–381 (Plenum, New York, 1985).
- Jungers, W. L. *Yb. phys. Anthropol.* **27**, 73–97 (1984).
- Preuschoft, H. in *Gravity, Posture and Locomotion in Primates* (eds Jouffroy, F. K., Stack, M. H. & Niemitz, C.) 109–127 (Il Sedicesimo, Firenze, 1990).
- Jungers, W. L. in *The Lesser Apes: Evolutionary and Behavioural Biology* (eds Preuschoft, H., Brockelman, W. & Creel, N.) 146–169 (Edinburgh Univ. Press, 1984).
- Begun, D. R. *Science* **257**, 1929–1933 (1992).
- Begun, D. R. *Yb. phys. Anthropol.* **37**, 11–63 (1994).

- Pilbeam, D., Rose, M. D., Barry, J. C. & Ibrahim Shah, S. M. *Nature* **348**, 237–239 (1990).
- Pilbeam, D. & Simons, E. *Nature* **229**, 408–409 (1971).
- Kelley, J. & Pilbeam, D. in *Comparative Primate Biology* (eds Swindler, D. R. & Erwin, J.) 316–411 (Liss, New York, 1986).
- Ruff, C. B., Walker, A. & Teaford, M. F. *J. hum. Evol.* **18**, 515–536 (1989).
- Ruff, C. B. in *Body Size in Mammalian Paleobiology. Estimation and Biological Implications* (eds Damuth, J. & MacFadden, B. J.) 119–149 (Cambridge Univ. Press, 1990).
- McHenry, E. M. *Am. J. phys. Anthropol.* **87**, 407–431 (1992).

ACKNOWLEDGEMENTS. We acknowledge helpful comments and discussion from P. Andrews, R. D. Martin, M. Pickford, D. Pilbeam, L. Alcalá and J. Agustí. We thank the following people for permission to see material of extant primates in their care: J. Barreiro from the Museo Nacional de Ciencias Naturales de Madrid; B. Engesser from the Naturhistorisches Museum Basel; S. Filella from the Zoological Garden of Barcelona; J. Franzen and G. Storch from the Senckenberg Museum Frankfurt; H. Schliemann from the Zoologisches Institut der Universität Hamburg; P. Schmid from the Anthropologisches Institut der Universität Zürich; and F. Unbe from the Museu de Zoologia de Barcelona. We thank O. Clavell (Museu Arqueologic de Barcelona) for the photographs and A. Galobart and I. Pellejero for restoring the specimens. The excavations were funded by the Diputació de Barcelona, the Generalitat de Catalunya, the CAYCIT and the Wenner Gren Foundation.

Detection and cellular localization of elemental sulphur in disease-resistant genotypes of *Theobroma cacao*

Richard M. Cooper^{*†}, Mario L. V. Resende^{*}, Julie Flood^{*}, Michael G. Rowan[‡], Michael H. Beale[§] & Ursula Potter^{||}

^{*} School of Biology and Biochemistry, [‡] School of Pharmacy and Pharmacology, and ^{||} School of Materials Science, University of Bath, Bath, Avon BA2 7AY, UK

[§] IACR, Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Avon BS18 9AF, UK

DISEASE-RESISTANT genotypes are the basis for controlling many major microbial pathogens of economic plants. Resistance is often linked with organic, antimicrobial phytoalexins, produced *de novo* in cells surrounding apoptotic or 'hypersensitive' cells that die rapidly after contact with incompatible pathogens^{1–3}. Here we report the production by resistant genotypes of *Theobroma cacao* of four phytoalexins in response to a xylem-invading fungal pathogen; these comprised two phenolics, a triterpenoid and, highly unusually in a higher eukaryote, elemental sulphur as cyclooctasulphur S₈. Energy-dispersive X-ray microanalysis revealed a high accumulation of sulphur only in cells and structures in potential contact with the vascular pathogen; that is, xylem parenchyma, xylem vessel cell walls and gels occluding vessels. Our data provide a rare example of cellular localization of an antimicrobial substance and evidence for the first time for accumulation of elemental sulphur in a plant linked with a resistance response; this discovery comes centuries after man first used elemental sulphur as a potent fungicide^{4,5}.

World production of cocoa lost to diseases is estimated at 20–30% per annum⁶ and in Brazil alone, a centre of origin of *Theobroma cacao*, there are ~22,000 accessions available as possible sources of resistance⁷. As part of a programme to improve techniques for rapid testing of new genotypes of *T. cacao* for resistance to the devastating soil-borne vascular pathogen *Verticillium dahliae*⁷, we investigated preformed and induced antimicrobial compounds as possible chemical indicators of disease resistance.

Resistance of two genotypes of *V. dahliae* was linked with the production, starting between 3 and 10 days after inoculation, of four antifungal compounds. These were isolated by chromatography and their structures determined by NMR and mass spectrometry (Fig. 1). The most abundant, and also most polar, antifungal compound (Table 1) was the pentacyclic triterpene,

[†]To whom correspondence should be addressed.

arjunolic acid ($2\alpha,3\beta,23$ -trihydroxyolean-12-en-28-oic acid) (**1**), the identity of which was confirmed by comparison by gas chromatography-mass spectroscopy (GC-MS) of its methyl ester with an authentic sample⁸. Of intermediate polarity and of least abundance were the related phenolics 3,4-dihydroxyacetophenone (**2**) and 4-hydroxyacetophenone (**3**). The least polar compound was unambiguously identified as elemental sulphur (**4**) by GC-MS (S_8 breaking down successively to S_2) and X-ray crystallography. It was initially suspected to be an artefact of extraction from a sulphur-rich organic compound. However, there was a lack of 1H or ^{13}C NMR signals that would result from a sulphur-containing organic compound that undergoes thermolysis on GC-MS to give free S. Also, sulphur was consistently extracted with a wide polarity range of solvents and then detected as elemental sulphur by thin-layer chromatography (TLC; ref. 9) without further purification required for GC-MS. Sulphur-containing phytoalexins are known only in crucifers¹⁰, and these comprise few sulphur atoms per molecule. Its biosynthetic origin is uncertain; inorganic sulphate taken up by higher plants is

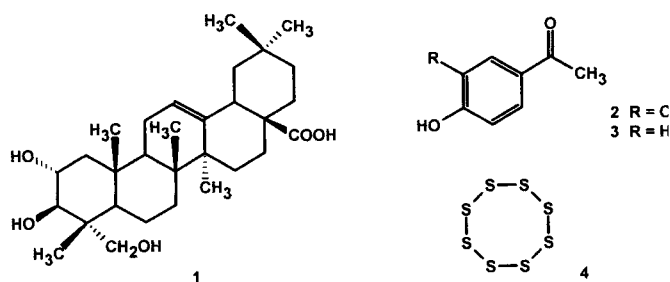


FIG. 1 Structures of antifungal compounds isolated from stems of resistant cocoa infected with *V. dahliae*.

METHODS. Cocoa seedlings (100 days old) of resistant genotypes Pound-7 or West African Amelonado or susceptible ICS-1 were stem-puncture-inoculated with *V. dahliae* conidia ($\sim 5 \times 10^6$ in 20 punctures per stem⁷). At intervals, sections of the wood (minus bark and pith) were comminuted in liquid N_2 in a mill, then extracted with diethyl ether in a Soxhlet apparatus for 48 h. The recovered extract was fractionated by flash chromatography on silica with diethyl ether/petroleum ether/methanol mixtures. Fractions were combined on the basis of TLC analysis, locating antifungal compounds by spraying with *V. dahliae* conidia in nutrient solution and incubating them for 48 h at 25°C; dark (melanized) mycelium formed on the silica except for zones of inhibition, which appeared white. The four compounds identified by this process were further purified to homogeneity by flash chromatography and preparative TLC. NMR/MS data: **1**, arjunolic acid 1H NMR (400 MHz CD_3OD) δ : 5.25 (t, $J = 3.5$ Hz, H-12), 3.68 (ddd, $J = 11, 9.5$ and 4.5 Hz, H-2 β), 3.49 and 3.26 (2d, $J = 11$ Hz, H-23), 3.35 (d, $J = 9.5$, H-3 α), 1.17, 1.02, 0.94, 0.91, 0.81 and 0.69 (6s, $6 \times CH_3$), ^{13}C NMR (100 MHz, CD_3OD) δ : 181.8 (s, C-28), 145.4 (s, C-13), 123.4 (d, C-12), 78.3 (d, C-3), 69.7 (d, C-2), 66.5 (t, C-23), 48.9 (d, C-5), 48.3 (d, C-9), 47.84 (t, C-1), 47.6 (s, C-17), 47.2 (t, C-19), 44.1 (s, C-4), 43.0 (s, C-14), 42.7 (d, C-18), 40.6 (s, C-8), 39.0 (s, C-10), 34.9 (t, C-21), 33.8 (t, C-7), 33.6 (q, C-29), 33.3 (t, C-22), 31.69 (s, C-20), 28.8 (t, C-15), 26.5 (q, C-27), 24.6 (t, C-16), 24.0 (t, C-11), 24.0 (q, C-30), 19.1 (t, C-6), 17.8 (q, C-26), 17.6 (q, C-25) and 13.8 (q, C-24). The methyl ester, prepared with diazomethane, was identical to an authentic sample⁸ when compared by GC-MS as the trimethylsilyl derivative. Kovats Retention Index, 3806. MS m/z (rel.int.) 718 (M^+ 0.2%), 658(0.3), 628(1.9), 613(1.7), 538(7), 276(7), 262(85), 231(25), 203(96), 147(38), 75(38) and 73(100). **2**, 3,4-dihydroxyacetophenone 1H NMR (270 MHz, CD_3OD) δ : 7.55 (m, H-2, H-6), 6.92 (d, $J = 8$ Hz, H-5), 2.6 (s, CH_3) MS m/z 152(M^+ , 53%), 137(100) and 109(30). **3**, 4-hydroxyacetophenone 1H NMR (270 MHz, $CDCl_3$) δ : 7.91(d, $J = 9$ Hz, H-2, 6), 6.90(d, $J = 9$ Hz, H-3, 5), 2.59(s, CH_3) MS m/z 136(M^+ , 36%), 121(100), 93(31). **4**, Sulphur MS m/z 256 (S_8 , 15%), 224(9), 192(30), 160(28), 128(21), 96(20) and 64(100). X-ray crystallography showed these crystals to be identical to rhombic sulphur. After extractions with hexane, diethyl ether, $(CH_3)_2CO$, CS_2 or methanol the antifungal compound always co-chromatographed with elemental sulphur by TLC on silica gel run with diethyl ether/petroleum ether/(methanol) (6:3:1); efficacy of extraction decreased with increasing polarity.

TABLE 1 Toxicities and levels *in vivo* of four antifungal substances from stems of a resistant cocoa genotype inoculated with *V. dahliae*

Compound	ED ₅₀ ($\mu g\ ml^{-1}$) germination*	Concentration in stems $\mu g\ g^{-1}$ f.wt†
Arjunolic acid	12.8	168
3,4-Dihydroxyacetophenone	92.5	4.9
4-Hydroxyacetophenone	7.2	2
Sulphur	3.6	51(116)‡

* Toxicity of pure compounds was determined against germination of *V. dahliae* conidia⁷ in distilled water at pH 7.0 on glass slides at 25°C for 15 h; from assessment of 100 conidia on each of four replicate slides ED₅₀ (estimated concentration of compound which reduced germination by 50%) was estimated by linear regression analyses.

† Levels of each compound *in vivo* are conservative figures based on final yield of pure compound 15 days after inoculation of genotype Pound-7. No significant toxicity was detected in extracts from the susceptible genotype ICS-1.

‡ Sulphur continued to accumulate up to 60 d and in some experiments with genotype West African Amelonado reached $\sim 116\ \mu g\ g^{-1}$ fresh weight according to TLC assay⁹. No sulphur was detected in mycelium of *V. dahliae* or from cell-free culture medium.

usually converted to organic sulphur compounds such as amino acids¹¹. Elemental sulphur formation is a property of many prokaryotes such as photosynthetic and chemoautotrophic sulphur bacteria or cyanobacteria. In eukaryotes, production of S_8 has been described for only a few fungi and algae¹²⁻¹⁴. Free sulphur is the antibiotic component of the red algae *Ceramium rubrum* and *Erythrophyllum delesserioides*¹³. However, other evidence has recently come to light concerning alternative metabolism of sulphur in plants. Elemental sulphur (S_8) of plant origin has been found in epicuticular wax of several gymnosperms and angiosperms by Kylin *et al.*¹⁵ They speculated on a role in constitutive defence against fungi. Also, elemental sulphur (S_8) was produced by suspensions of illuminated spinach chloroplasts supplied *in vitro* with $^{35}SO_4^{2-}$ (ref. 12).

To implicate an antimicrobial compound in disease resistance it must be shown to be produced in sufficient quantity, in the right place and at the right time to inhibit the pathogen; these criteria have rarely been met. At first sight, only sulphur and arjunolic acid were present at fungitoxic levels in extracts from infected wood of resistant cocoa stems. Sulphur afforded the unusual opportunity of attempting cellular localization of an antifungal substance by coupled scanning electron microscopy and energy-dispersive X-ray microanalysis (SEM-EDX). Previous localization of phytoalexins has been at the level of light microscopy by fluorescence and ultraviolet microphotometry^{1,16,17}, by histochemistry¹⁸ and by laser microprobe mass analysis¹⁹, and quantification *in situ* has been achieved by radioimmunoassay²⁰. In this study, high concentrations of sulphur were found only in scattered xylem parenchyma cells in direct contact with xylem vessels, within vessel walls and in gels (synonymous with gums) occluding xylem vessels (Fig. 2). These specific locations are of direct relevance to resistance to xylem-invading *V. dahliae*. The pathogen is often in intimate contact with xylem vessel walls²¹ and must penetrate them when a vessel end wall is reached by hyphae or by propagules carried in the transpiration stream²². Vascular gels comprise mainly polysaccharides, are induced by infection and seem to trap the upward progress of vascular parasites in resistant genotypes²¹.

These structures can be breached enzymatically by the pathogen²² unless stabilized by phenolic crosslinking or impregnated with antimicrobial compounds. Thus impregnation of host cell walls and gels by sulphur should provide an effective defence against systemic invasion. Analogously, in cotton infected with *V. dahliae*, terpenoid phytoalexins are formed in xylem parenchyma cells, exuded into vessels and absorbed by cell walls, the pathogen and vascular occlusions¹⁸.

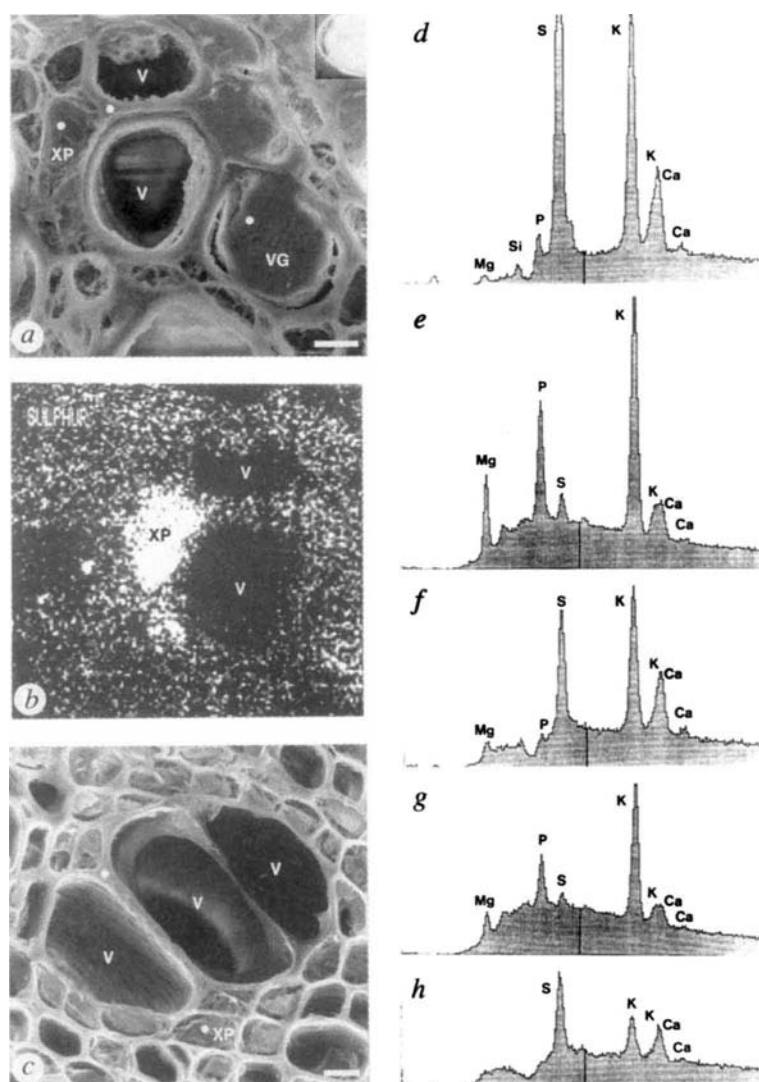


FIG. 2 Distribution of sulphur in vascular tissues of inoculated and control stems of *T. cacao*. XP, xylem parenchyma cell; V, xylem vessel; VG, vascular gel. The X-ray map (b) was produced from the equivalent area in a; spot analyses (d–h) were made from a and c at the points marked *; the bar on a, b represents 10 μ m. In uninoculated stems (c) localized accumulation of S was never observed and spot analyses of XP cells revealed S as a minor peak with K the predominant element (e). Inoculated stems (a) contained scattered XP cells with a high accumulation of S evident from the X-ray dot maps (b); X-ray spectra revealed S peaks often equivalent to or greater than K (d). Adjacent xylem vessel cell walls also contained an unusually high proportion of S (f) compared with walls in control plants (g). In vascular gels, S was sometimes the main element (h); the non-cytoplasmic nature of these structures is apparent from the low K content.

METHODS. EDX was performed on transverse sections of lyophilized stems coated with carbon and viewed in a Jeol 6310 scanning electron microscope. General areas covering many cells or subcellular spot analyses of elements were made with an Oxford Instruments X-ray analyser. The data are representative of 54 spot analyses performed on 10 areas of vascular and ground tissues from upper regions of stems from infected and uninfected plants, inoculated and sampled on two occasions.

No elemental sulphur was recovered from intact or wounded plants or from the susceptible genotype. The accumulation of sulphur only in *V. dahliae*-infected stems of resistant genotypes might imply specific elicitation, but more probably it reflects accumulation in cells lacking metabolic capabilities, in other words dead xylem vessels and xylem parenchyma cells that have responded hypersensitively to the pathogen. Phytoalexins such as rishitin (from potato) and glyceollin (from soybean) are known to be produced, but do not accumulate in wounded but uninoculated tissues of these plants because they are metabolized by living plant cells²³; this rapid turnover may reflect the non-specific toxicity of many phytoalexins². The death of scattered xylem parenchyma cells is typical for vascular diseases and results in the characteristic discoloration of xylem by oxidized phenolics^{18,21}. Phytoalexins often accumulate to very high levels in such localized necrotic host cells after production in adjacent living cells^{2,3}.

Plant cells or chloroplasts such as those from wheat and spinach can metabolize S_8 (refs 5, 12, 24), but in *T. cacao* stems its persistence over at least 60 days further suggests that S_8 is unavailable to living cells. The persistence of sulphur also suggests that levels are fungitoxic because many fungi can also metabolize S_8 , but only if levels are sublethal²⁵. Indeed, S_8 levels at sites colonized by *V. dahliae* are likely to be highly fungitoxic, on the basis of the high amounts obtained from extracts of stems of which only a very small proportion of cells, according to SEM–EDX, contained accumulated sulphur.

Most plant species produce several chemically related phyto-

alexins². However, these inducible antifungal substances, which have not previously been found in *T. cacao*, result from three unrelated biosynthetic or metabolic pathways, although the three organic compounds have been reported as antimicrobials from other plant species. Arjunolic acid is found in many tropical and fruit trees and its derivatives affect some viral and bacterial animal pathogens and show inhibitory effects on skin-tumour promoters²⁶; this is the first report of toxicity to a plant pathogen. 4-Hydroxyacetophenone is a constitutive antifungal compound in gymnosperms²⁷ and an indicator of general stress in many species within the Pinaceae²⁸. A derivative of the dihydroxy analogue may function as a phytoalexin in roots of *Sanguisorba minor* (Rosaceae)²⁹.

Other than strengthening or crosslinking of plant cell walls by calcium or silicon³⁰, inorganic elements have not previously been implicated directly in disease resistance. Other recent work¹⁵ and this research shows that diverse plant species can produce and accumulate elemental sulphur in cellular components such as epicuticular wax, cell walls and vascular gels, structures that can play key roles in defence³⁰. In view of its potent fungicidal and acaricidal activities it is tenable that man's probably oldest pesticide^{4,5} already functions in this role in living plants. □

Received 28 August; accepted 2 November 1995.

1. Mansfield, J. M. in *Natural Antimicrobial Systems* (eds Gould, G. W. et al.) 191–205 (Bath Univ. Press, 1986).
2. Bailey, J. A. & Mansfield, J. W. *Phytoalexins* (Blackie, Glasgow, 1982).

3. Darvill, A. G. & Albersheim, P. A. *Rev. Pl. Physiol.* **35**, 243–275 (1984).
4. Tweedy, B. G. *Residue Rev.* **78**, 43–68 (1981).
5. Legris-Delaporte, S., Ferron, F., Landry, J. & Coster, C. *Pl. Physiol.* **85**, 1026–1030 (1987).
6. Holderness, M. *Biochemist* **2**, 11–13 (1993).
7. Resende, M. L. V., Flood, J. & Cooper, R. M. *Pl. Path.* **44**, 374–383 (1995).
8. Yaguchi, Y., Sakurai, N., Nagai, M. & Inoue, T. *Chem. Pharm. Bull.* **36**, 1419–1424 (1988).
9. Davies, J. R. & Thuringham, S. T. J. *Chromatogr.* **35**, 513–518 (1968).
10. Devys, M. *et al. Tetrahedron Lett.* **29**, 6447–6448 (1988).
11. Young, L. & Maw, G. A. *The Metabolism of Sulphur Compounds* (Methuen, London, 1958).
12. Joyard, J., Fores, E., Blée, E. & Douce, R. *Pl. Physiol.* **88**, 961–964 (1988).
13. Izak, R. R., Stierle, D. B. & Sims, J. J. *Phytochemistry* **21**, 229 (1982).
14. Steudel, R., Holdt, G., Visscher, P. T. & Van Gernerden, H. *Arch. Microbiol.* **153**, 432–437 (1990).
15. Kylin, H., Atuma, S., Hovander, L. & Jensen, S. *Experientia* **50**, 80–85 (1994).
16. Mansfield, J. W., Hargreaves, J. A. & Boyle, F. C. *Nature* **252**, 316–317 (1974).
17. Mayana, S. & Tani, T. *Physiol. Pl. Path.* **21**, 141–149 (1982).
18. Mace, M. E., Bell, A. A. & Beckman, C. H. *Can. J. Bot.* **54**, 2095–2099 (1976).
19. Moesta, P., Seydel, U., Lindner, B. & Grisebach, H. *Z. Naturf.* **37c**, 748–751 (1982).
20. Hahn, M. G., Bonhoff, A. & Grisebach, H. *Pl. Physiol.* **77**, 591–601 (1985).
21. Cooper, R. M. in *Plant Disease Control: Resistance and Susceptibility* (eds Staples, R. C. & Toennies, G. H.) 105–142 (Wiley, New York, 1981).
22. Durand, P. K. & Cooper, R. M. *Physiol. molec. Pl. Path.* **32**, 343–362 (1988).
23. Ward, E. W. B., Stoessl, A. & Stothers, J. B. *Phytochemistry* **16**, 2024–2025 (1977).
24. Jolivet, P., Bergeron, E., Zimierski, A. & Meunier, J.-C. *Phytochemistry* **38**, 9–14 (1995).
25. Beffa, T. *Can. J. Microbiol.* **39**, 736–741 (1993).
26. Diallo, B., *et al. J. nat. Prod.* **52**, 879–881 (1989).
27. Gottstein, D. & Cross, D. *Trees Struct. Funct.* **6**, 55–68 (1992).
28. Jensen, J. S. & Lokke, H. J. *Pl. Dis. Prof.* **97**, 328–338 (1990).
29. Kokubun, T., Harborne, J. B. & Eagles, J. *Phytochemistry* **35**, 331–333 (1994).
30. Ride, J. P. in *Biochemical Plant Pathology* (ed. Callow, J.) 215–236 (Wiley, Chichester, 1983).

ACKNOWLEDGEMENTS. We thank K. Molloy (School of Chemistry, University of Bath) for X-ray crystallography, P. Gaskin (IACR, Long Ashton) for carrying out the GC-MS analysis and Nobuko Sakurai (Faculty of Pharmaceutical Sciences, Hoshi University, Tokyo) for an authentic sample of arjunolic acid.

Patterning of the *Drosophila* embryo by a homeodomain-deleted Ftz polypeptide

John W. R. Copeland, Andrzej Nasiadka,
Bruce H. Dietrich & Henry M. Krause*

Banting and Best Department of Medical Research, University of Toronto, Charles H. Best Institute, and Department of Medical and Molecular Genetics, 112 College Street, Toronto, Ontario M5G 1L6, Canada

*To whom correspondence should be addressed

HOMEODOMAIN proteins regulate diverse developmental processes in a wide range of organisms, yet bind *in vitro* to DNA sequences that are remarkably similar¹. This has raised the fundamental question of how target gene specificity is achieved *in vivo*. The *Drosophila fushi tarazu* protein (Ftz) contains a homeodomain² and is required for the formation of alternate segments³. We have shown previously that a homeodomain-deleted Ftz polypeptide (FtzΔHD), incapable of binding DNA *in vitro*, could regulate endogenous *ftz* gene expression⁴. Here we test FtzΔHD activities in a *ftz* mutant background and find that, surprisingly, FtzΔHD

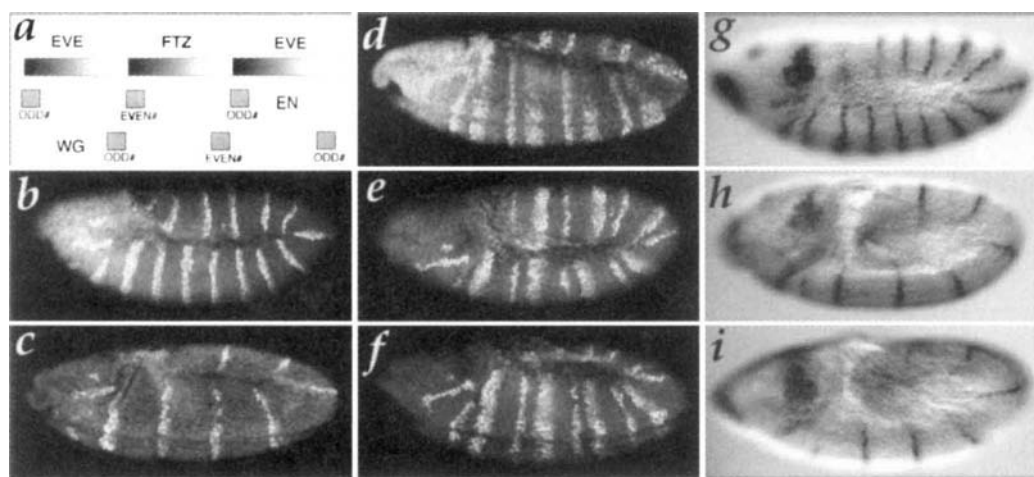
can directly regulate *ftz*-dependent segmentation, suggesting that it can control target gene expression through interactions with other proteins. A likely candidate is the pair-rule protein Paired (Prd). FtzΔHD bound directly to Prd *in vitro* and required Prd to repress *wingless* *in vivo*. These results emphasize the pivotal importance of protein–protein interactions in homeodomain protein function.

Ftz polypeptides with severe homeodomain mutations fail to accumulate because of an inability to autoregulate at low levels^{5,6}. Accordingly, FtzΔHD (deleted for most of helices 1 and 3 and all of helix 2), expressed under control of the *ftz* gene promoter, does not rescue *ftz*-dependent segments⁵. We bypassed the need for autoregulation by expressing 'autoregulated' levels of Ftz and FtzΔHD, in *ftz*[−] embryos, using the *hsp70* promoter. We then asked if FtzΔHD could directly regulate other *ftz*-dependent processes.

We looked first at expression of the segment polarity genes *engrailed* (*en*), and *wingless* (*wg*), which are positively⁷ and negatively⁸ regulated by *ftz*. Figure 1(b–f) shows that, in a *ftz* mutant background, Ftz and FtzΔHD could both rescue *ftz*-dependent *en* stripes. Both polypeptides were also capable of repressing alternate *wg* stripes in the *ftz*[−] background (Fig. 1g–i). Thus, transiently expressed Ftz and FtzΔHD could both regulate *ftz*-dependent genes in the absence of an endogenous *ftz* gene.

The most comprehensive measure of Ftz activities is the production of *ftz*-dependent segments. *ftz*[−] embryos generate a

FIG. 1 Rescue of *En* expression (b–f) and repression of *wg* (g–i) in *ftz*[−] embryos. a, Relative expression patterns of *ftz*, *eve*, *en* and *wg* over a span of three segments. Intensity gradients reflect the narrowing of stripes during gastrulation. b–f, Expression of *en* in HSF; *ftz*^{w20} and HSFΔHD; *ftz*^{w20} embryos. Embryos were double-stained for *En* and *hunchback* promoter-driven β -galactosidase expression. β -Gal expression in the head (b, d) indicates a *ftz*⁺ genetic background. b, Non-heat-shocked, HSF; *ftz*⁺ embryo; 14 evenly spaced *En* stripes are visible. c, Non-heat-shocked HSF; *ftz*[−] embryo; even-numbered *En* stripes are missing. d, Heat-shocked HSF; *ftz*[−] embryo; even numbered stripes have expanded anteriorly. e, Heat-shocked HSF; *ftz*[−] embryo; missing *En* stripes have been rescued and are expanded as in d. f, Heat-shocked HSFΔHD; *ftz*[−] embryo; even numbered *En* stripes are rescued and expanded as in d and e. g–i, *wg* mRNA expression in g, wild-type; h, HSF; *ftz*[−]; and i, HSFΔHD; *ftz*[−] embryos. Both Ftz and FtzΔHD were able to repress odd-numbered *wg* stripes, as previously shown in *ftz*⁺ embryos⁴. METHODS. HSF and HSFΔHD fly lines have previously described⁴. HSF and HSFΔHD constructs, homozygous on either the



1st or 2nd chromosomes, were introduced into a *ftz*^{w20} (or *ftz*^{9H34}, not shown) protein-null background, generating the stocks HSF; *Ki*, *ftz*^{w20}/TM3 and HSFΔHD; *Ki*, *ftz*^{w20}/TM3. The TM3 balancer used carries a *hunchback-lacZ* reporter¹⁹. Embryos were heat shocked for 8–10 min as previously described²⁰. Embryos were fixed following a 1.5-h recovery and stained, either for *En* and β -gal protein (monoclonal antibodies obtained from S. Coté and BioRad, respectively), or *wg* mRNA as previously described¹⁹.

pair-rule phenotype as shown in Fig. 2*b*. Addition of low levels of exogenous Ftz or FtzΔHD, partially replaced the missing segments and partially deleted the intervening segments (Fig. 2*c,d*). At higher levels, both polypeptides were able to rescue the *ftz*-dependent segments fully, and to delete the intervening segments (Fig. 2*e,f*). The same 'anti-*ftz*' phenotype is generated in a *ftz*⁺ background^{4,9,10}. Cellular levels of expression required for these effects were equivalent to levels expressed in wild-type embryos

(Fig. 2 legend). Thus, a short pulse of exogenous Ftz or FtzΔHD expression was sufficient to direct *ftz*-dependent segmentation.

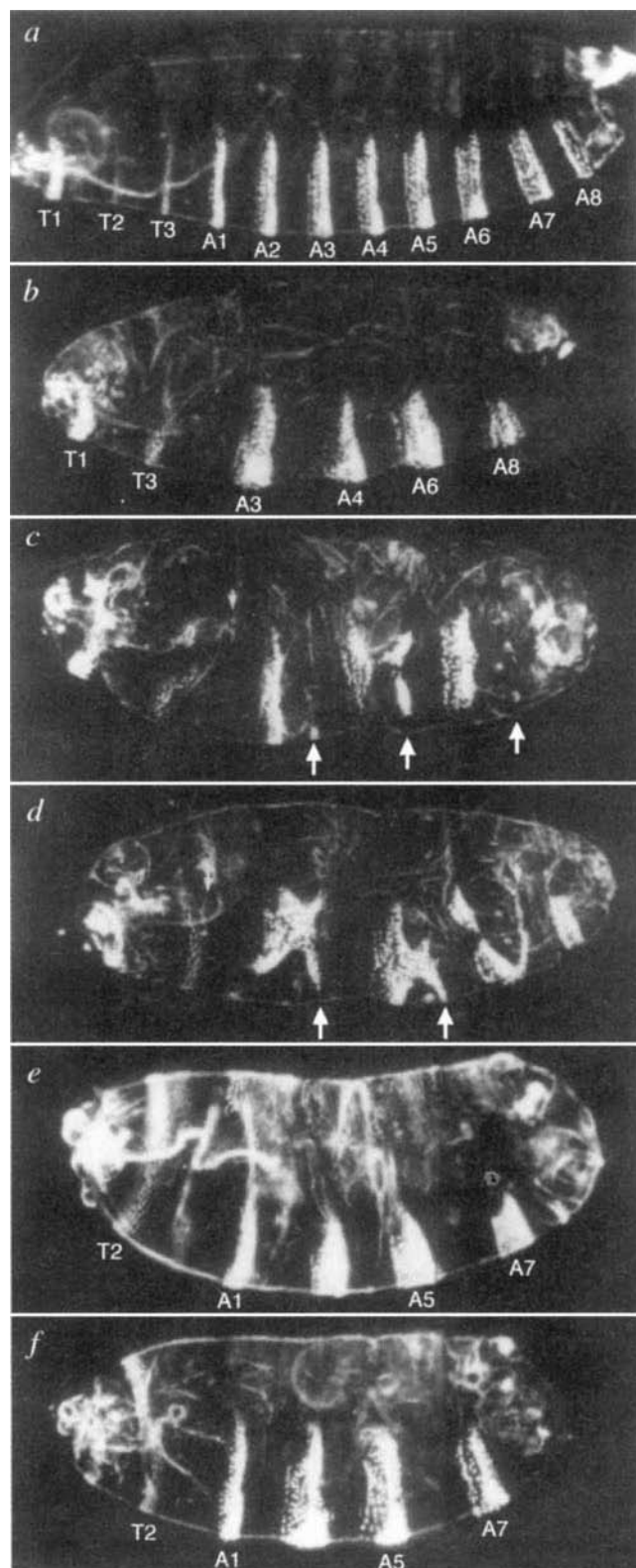
These results indicate that the DNA-binding activity of the highly conserved Ftz homeodomain is dispensable for many or most Ftz-dependent pair-rule functions, so long as sufficient levels of protein are expressed. This interpretation agrees with recent results of I. Duncan and co-workers, who have observed that the homeodomain of the endogenous *ftz* gene can be deleted, so long as the remaining regions contain secondary mutations that render the protein more stable (personal communication).

The ability of FtzΔHD to regulate *ftz*-dependent segmentation, despite its inability to bind DNA, suggests that it can interact with target gene promoters by contacting other DNA-bound proteins. Several criteria make the pair-rule protein Paired (Prd) a prime candidate for such a protein; the expression domains of Ftz and Prd are largely overlapping¹¹, the two genes interact genetically to position stripes of *wg*⁸ and *gooseberry*^{12,13}, and Prd binds DNA, possessing both homeodomain¹⁴ and Prd domain^{15,16} DNA-binding motifs.

Prd was expressed in reticulocyte lysates and passed over Ftz affinity columns to test for a direct Ftz-Prd interaction (Fig. 3). Prd was specifically retained on the column, whereas all other proteins passed through. Likewise, Fig. 3*c* shows that Ftz expressed using either *in vitro* or baculovirus expression systems bound to Prd columns as expected. These interactions were likely to be direct because pure baculovirus-expressed Ftz bound to columns of purified Prd, or far-western blots of Prd (not shown), in the presence or absence of competitors. Importantly, FtzΔHD (Fig. 3*c*) bound the Prd columns with equal efficiency. In contrast, deletion of amino acids 101–150 (FtzΔN) dramatically reduced retention on the column (Fig. 3*c*).

In the embryo, genetics and relative expression patterns (Fig. 4*a*) make *wg* the best candidate for a gene that is coordinately regulated by Ftz and Prd^{8,11}. Prd overlaps all 14 *wg* stripes⁸, and in the absence of Prd, initiation of all 14 stripes is severely compromised (Fig. 4*c*). When Prd was ubiquitously expressed using the *hsp70* promoter, all 14 *wg* stripes widened immediately after heat shock (Fig. 4*d*). These results suggest that Prd may be a direct activator of *wg*.

Conversely, Ftz may be a direct repressor of *wg*. Initiation of *wg* stripes in *ftz*-dependent parasegments begins when Ftz stripes retract⁸. In *ftz*[−] embryos, *wg* stripes initiated earlier, filling the entire *ftz*-dependent parasegments (Fig. 4*e*), and coinciding with expression of Prd¹⁷. These broad stripes were rapidly repressed following low levels of Ftz induction (Fig. 4*f*), and in an even less



◀ FIG. 2 Rescue of *ftz*-dependent segments. HSF; *ftz*[−] (*a,b,c,e*) and HSFΔHD; *ftz*[−] (*d,f*) embryos were heat-shocked for either 0 (*a,b*), 5 (*c,d*) or 10 min (*e,f*) and then allowed to develop for 24 h before preparation of cuticles. *a*, 74% (*n* = 180) of non-heat-shocked embryos generated wild-type cuticles. *b*, About 23% had alternate segments deleted. Few (3–5%) *ftz* mutant cuticles remained after a 5-min heat-shock. About 70% of cuticles were near wild-type with most of the rest resembling *c* for HSF; *ftz*[−] progeny (20%, *n* = 174) and *e* for HSFΔHD; *ftz*[−] progeny (22%, *n* = 204). After a 10-min heat-shock, 91% of HSF; *ftz*[−] progeny (*n* = 156) resembled the anti-*ftz* cuticle pattern shown in *d*, whereas 85% of HSFΔHD; *ftz*[−] progeny (*n* = 287) resembled the anti-*ftz* cuticle pattern shown in *f*. Remaining patterns were an assortment of wild-type and mutant patterns also seen in heat-shocked and non-heat-shocked controls. METHODS. Collection of 20-min embryos were made, heat-shocked and staged as previously described²⁰. Results from three experiments were tabulated. Similar phenotypic distributions were observed using a marked HSF line (*hsf2*; *ftz*^{w20}, *Df109*, *Ki*/TM1, red; obtained from P. Lawrence) and a marked HSFΔHD line (*y*[−]; HSFΔHD; *ftz*^{w20}, *Ki*/TM3, *y*[−]; D. Hyduk and A. Percival-Smith, personal communication). The *hsf2* line of Struhl⁹, required shorter heat-shocks for rescue (5–6 min), because of more efficient Ftz induction. Levels of Ftz expression were determined by western blot analysis as previously described²¹, except that signals were chemiluminescent²² and quantified on a BioRad phosphorimager. Levels of Ftz and FtzΔHD expression were similar, and 10-min inductions yielded cellular levels of expression equivalent to those of endogenous Ftz.

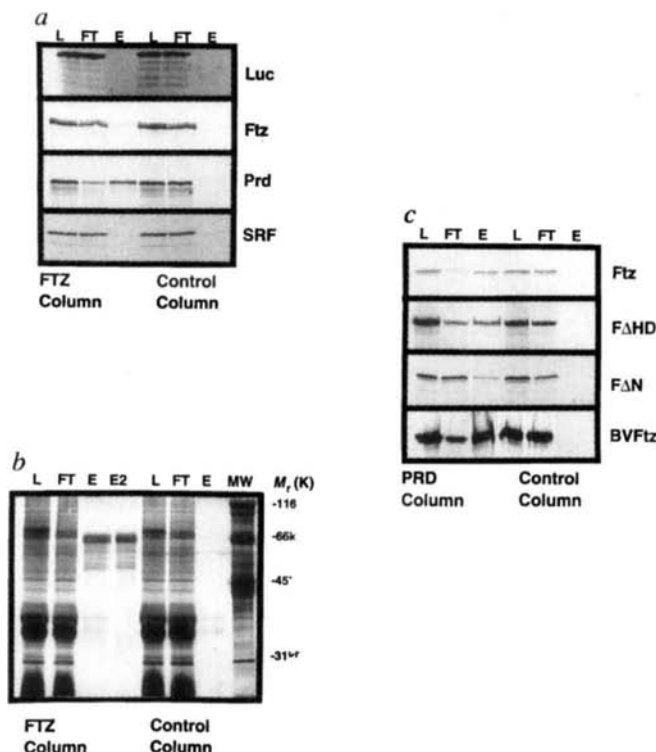


FIG. 3a, Ftz affinity chromatography. Reticulocyte lysate-translated proteins were passed over Ftz affinity columns or control columns. Equivalent amounts of the ^{35}S -labelled load (L), flow-through (FT), and eluate (E) fractions were subjected to SDS-PAGE and autoradiography. Prd was the only protein retained on the column (60–90%). Luciferase (Luc), Ftz and serum response factor (SRF) did not bind. b, Silver staining. Chromatography fractions shown in a, were silver-stained after SDS-PAGE. Most proteins present in the lysate flowed through the columns. Proteins in the Ftz column eluate correspond predominantly to Ftz stripped from the column. The same profile was eluted from a Ftz column loaded with buffer alone (E2). ^{35}S Prd migrates slower with a M_r of 66K. Other minor bands in the Ftz column eluate were present in the control column eluate. c, Prd affinity chromatography. *In vitro* translated Ftz, FtzΔHD, FtzΔN (amino acids 101–150), and purified baculovirus-expressed Ftz, were passed over Prd affinity columns and separated by SDS-PAGE. Bands were visualized by autoradiography or western blotting (BV Ftz). About 60–80% of Ftz and FtzΔHD polypeptides were retained on the columns. Retention of FtzΔN was nearly 10 times lower.

METHODS. Full-length *ftz* and *prd* cDNAs were subcloned into PET19B (Novagen) and expressed in *Escherichia coli*. Recombinant proteins were purified as described²¹, and affinity columns made by coupling the His-tagged proteins to Ni^{2+} resin as described by the manufacturer (Invitrogen) at a concentration of 1 mg ml^{-1} . Ni^{2+} resin alone served for the control columns. Columns were equilibrated in: 0.1M NaCl, 20 mM Tris pH 7.6, 1 mM EDTA, 10% Glycerol, 1% milk powder. *In vitro* translated protein (5 μl , Promega TNT kit) was diluted to 90 μl in column buffer, and passed over 50- μl columns. Full-length baculovirus-expressed Ftz (1 μg) purified to about 95% homogeneity (J.W.R.C., manuscript in preparation), was chromatographed as above, or using 1% casein or no competitor at all. Bound proteins were eluted with 2% SDS.

restricted fashion by FtzΔHD (Fig. 4g). In contrast, repression of *wg* was not observed when amino acids 101–150 (FtzΔN), required for Ftz to bind Prd *in vitro*, were removed (Fig. 4h). Similarly, when Prd was removed genetically, Ftz and FtzΔHD were unable to repress the residual, *prd*-independent *wg* stripes. These results suggest that recruitment of Ftz to the *wg* promoter requires a direct interaction with Prd, and that Ftz is then able to repress both Prd-dependent and Prd-independent *wg* expression (Fig. 4i, j).

The ability of Prd to recruit homeodomain-deleted Ftz polypeptides to DNA was previously suggested by cotransfection experiments¹⁸. However, in these studies an oligomerized piece

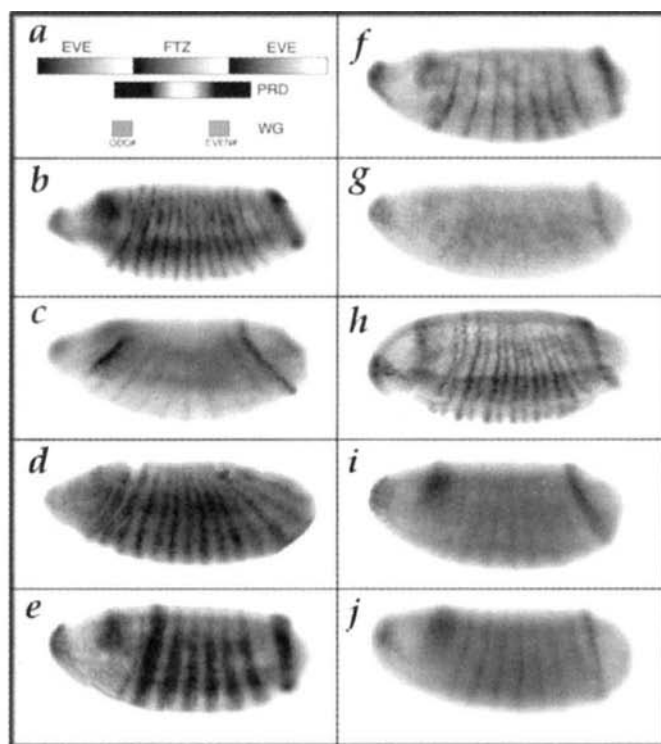


FIG. 4 Regulation of *wg* by Ftz and Prd. Whole-mount *in situ* hybridization was performed to monitor initiation of *wg* transcription (b–j). a, Summary of relative expression patterns of Ftz, Prd, Even-skipped (Eve) and *wg* over a 3-segment interval, as depicted in Fig. 1. b, Wild-type *wg* expression; 14 one-cell-wide stripes initiate simultaneously. c, Loss of *wg* expression in *prd*^{2.45} embryos; even-numbered stripes are missing, and odd-numbered stripes are about 3 times weaker than normal. All 14 *wg* stripes broadened anteriorly by 1–2 cells in HSPrd embryos within 15–20 min of a 10-min heat-shock (d). *wg* stripes initiated earlier in *ftz*⁻ embryos, filling the normal domains of *ftz* expression (e). Transcription of these broadened stripes was repressed within 15 min of a 6–8-min HSF (f) or HSFΔHD (g) induction. Note that FtzΔHD completely repressed all *wg* stripes whereas Ftz had more limited effects in the posterior-most regions of the Ftz expression domains. These even-numbered stripes recovered within an hour in both HSF and HSFΔHD embryos (Fig. 1h, i). h, No repression was observed in *ftz*⁺ embryos following a 10-min induction of FtzΔN. Phasing of stripes observed after longer recoveries (40 min shown) was a consequence of other retained activities (not shown). Ten-min inductions of Ftz (i) and FtzΔHD (j), in *prd*⁻ embryos, failed to repress residual (odd-numbered) *wg* stripes.

METHODS. Embryos were collected and heat-shocked as described in Fig. 1. *In situ* hybridization was performed as previously described²⁰. *prd*^{2.45} was used for i and j, and wild-type chromosomes were marked by the presence of an *evelacZ* reporter. The HSFΔN line was kindly provided by J. Ho and A. Percival-Smith. Images were captured using a Sony XC275 CCD camera and Northern Exposure software. Figures were compiled using Photoshop 3.0.

of the *en* promoter was used as the reporter, and the Ftz–Prd interaction resulted in synergistic activation of transcription. Thus, it appears that Ftz–Prd complexes may act as both repressors and activators of target gene expression. This differential activity probably depends on the placement of homeodomain and Prd domain binding sites within these promoters, or by the actions of promoter-specific auxiliary factors. In the embryo, FtzΔHD activated two of the three genes monitored (*en* and *ftz*). The requirement for Prd in these two processes has yet to be determined.

In summary, our results emphasize the importance of protein–protein interactions in homeodomain protein function. Ftz and Prd are both homeodomain proteins and, as a complex, are likely to have specificities and functions not exhibited individually. These may include both repression and activation of target gene promoters. Important questions now are how these effects are achieved and whether additional proteins are required. □

Received 5 September; accepted 27 October 1995.

- Hayashi, S. & Scott, M. P. *Cell* **63**, 883–894 (1990).
- Shepherd, J. C. W. et al. *Nature* **310**, 70–71 (1984).
- Wakimoto, B. T. & Kaufman, T. C. *Dev Biol.* **81**, 51–64 (1981).
- Fitzpatrick, V. D. et al. *Nature* **356**, 610–612 (1992).
- Furukubo-Tokunaga, K. et al. *Genes Dev.* **6**, 1082–1096 (1992).
- Schier, A. F. & Gehring, W. P. *Nature* **356**, 804–807 (1992).
- DiNardo, S. & O'Farrell, P. H. *Genes Dev.* **1**, 1212–1225 (1987).
- Ingham, P. W., Baker, N. E. & Martinez-Arias, A. A. *Nature* **331**, 73–75 (1988).
- Struhl, G. *Nature* **318**, 677–680 (1985).
- Ish-Horowitz, D. et al. *Cell* **57**, 223–232 (1989).
- Kilcherr, F. et al. *Nature* **321**, 493–499 (1986).
- Li, X., Gutjahr, T. & Noll, M. *EMBO J.* **12**, 1427–1436 (1993).
- Li, X. & Noll, M. *EMBO J.* **12**, 4499–4509 (1993).
- Frigerio, G. et al. *Cell* **47**, 735–746 (1986).
- Bopp, D. et al. *Cell* **47**, 1033–1049 (1986).
- Treisman, J. et al. *Genes Dev.* **5**, 594–604 (1991).
- Baumgartner, S. & Noll, M. *Mech. Dev.* **33**, 1–18 (1991).
- Ananthan, J. et al. *Molec. cell. Biol.* **13**, 1599–1609 (1993).
- Driever, W., Thoma, G. & Nüsslein-Volhard, C. *Nature* **340**, 363–367 (1989).
- Manoukian, A. & Krause, H. M. *Genes Dev.* **6**, 1740–1751 (1992).
- Krause, H. M., Klemenz, R. & Gehring, W. J. *Genes Dev.* **2**, 1021–1036 (1988).
- Gallagher, S. et al. in *Current Protocols in Molecular Biology* (eds Ausubel, F. A. et al.) 10.8.7–10.8.17 (Wiley, New York, 1993).

ACKNOWLEDGEMENTS. We thank A. Percival-Smith, P. Lawrence, I. Duncan and C. Desplan for freely exchanging reagents, unpublished results and ideas. En monoclonal antibody was kindly provided by S. Coté. The manuscript was improved by comments received from A. Spence, C. J. Ingles and P. Lawrence. This work was supported by grants from the Medical Research Council of Canada, and the National Cancer Institute of Canada. J.W.R.C. was supported in part by the Natural Sciences and Engineering Research Council of Canada and University of Toronto Open Fellowships.

Early computation of shape and reflectance in the visual system

Jennifer Sun & Pietro Perona

California Institute of Technology 116-81, Pasadena, California 91125, USA

A COMPELLING sense of three-dimensional shape may be conveyed by the photograph of an object. Cues such as contour, shading, perspective and occlusion, to name a few, contribute to this percept. Psychophysical experiments suggest that certain aspects of three-dimensional shape are computed rapidly and in parallel by the visual system^{1–7,15}. Here we report that reflectance is also computed rapidly; moreover, it is the apparent reflectance, rather than brightness or perceptual three-dimensional shape, that is the primary basis for discrimination during the early stages of visual processing.

In Fig. 1a, the background is composed of distractor patterns that are typically perceived as lit-from-above cubes. The target pattern, an upside-down version of the distractor pattern, pops out effectively. This pop-out is correlated with perceptual three-dimensional (3-D) shape^{3,5}, and is a parallel³ and fast (<80 ms)^{6,7} process. Similar results involving smoothly shaded patterns have also been reported^{1,2,5}. These results have been taken as evidence that mechanisms for computing 3-D shape feed into 'selection' mechanisms to give pop-out^{3,4}.

In experiment 1, displays were composed of simple patterns of two grey levels that may be perceived as shaded corners. This 3-D shape was imposed onto either the distractors (condition 1, shown in Fig. 2a) or the target (condition 2, shown in Fig. 2b) using stereo disparity cues and perspective projection. To isolate the effects of shape from shading, we constructed the same displays using unshaded wire-frame drawings (conditions 4 and 5, displays not shown).

We find that, for the shaded patterns, there is an asymmetry in performance between conditions 1 and 2 (Fig. 2c); it is easier to spot a flat target among 3-D distractors than to pick out a 3-D target among flat distractors. If discrimination were based directly on 3-D shape, the same asymmetry should occur for the wire-frame stimuli. However, we see from conditions 4 and 5 (Fig. 2d) that the asymmetry is in the opposite direction. In this case, it is easier to pick out a 3-D target among flat distractors than vice

versa. Moreover, for the wire-frame stimuli, discrimination is quite slow by comparison; display durations were 180 ms for the shaded conditions and 250 ms for the wire-frame ones. These findings are very difficult to reconcile with the view that fast discrimination is based directly on shape.

We suggest, as an alternative, that fast discrimination is based mainly on apparent reflectance. Perceptions of shape and reflectance are closely related (Fig. 1b). When a pattern is seen as flat, its luminance pattern is perceived as being due to changes in reflectance. But when a pattern is seen as having 3-D shape, as are the distractors in condition 1, the changes in its luminance pattern are discounted to some extent as shading due to shape, and its reflectance may be perceived as relatively constant. If apparent reflectance, rather than shape or brightness, is used as the feature for discrimination, a flat, multiple-reflectance target pattern in a background composed of single-reflectance 3-D shapes would be discriminated with ease (condition 1; Fig. 2a). Conversely, if the distractors were perceived as flat, the background would be perceived as having multiple reflectances. Against such a background, the target, whether 3-D with a single reflectance (condition 2; Fig. 2b) or 2-D with multiple reflectances (condition 3; display not shown), would be difficult to spot.

If our hypothesis is correct, 3-D shape could also be used to decrease the saliency of a target. In experiment 2, we used triptych patterns that are easy to discriminate when perceived as flat. Three-dimensional shape was imposed onto both the distractor and the target patterns (conditions 1 and 2 in Fig. 3a) using stereo disparity cues and perspective projection. Our hypothesis predicts that, when an appropriate 3-D shape is imposed, the differences in luminance of the vertical bands will be discounted, to some extent, as shading due to shape. Instead of flat patterns with distinctly

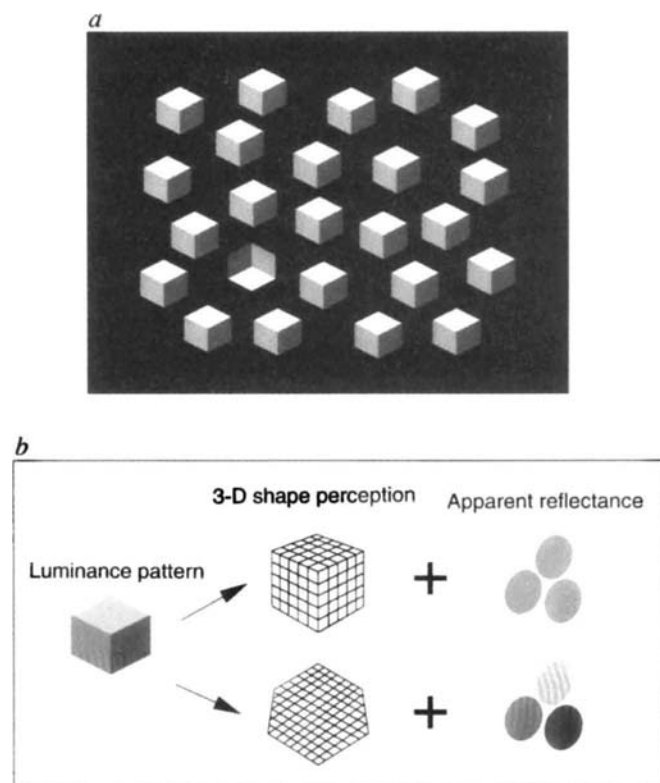


FIG. 1 a, Against a background of top-lit, convex cubes, the 'odd-man-out' target can be discriminated with ease^{3,6,7}. b, The perception of shaded patterns may be represented with a shape map and a reflectance map. If the shaded pattern leads to the perception of a 3-D shape, then the apparent differences in luminance are perceived as shading, and the resulting apparent reflectance is uniform. If the perception is of a flat surface, then the apparent luminance differences are attributed to changes in reflectance^{9–13}.

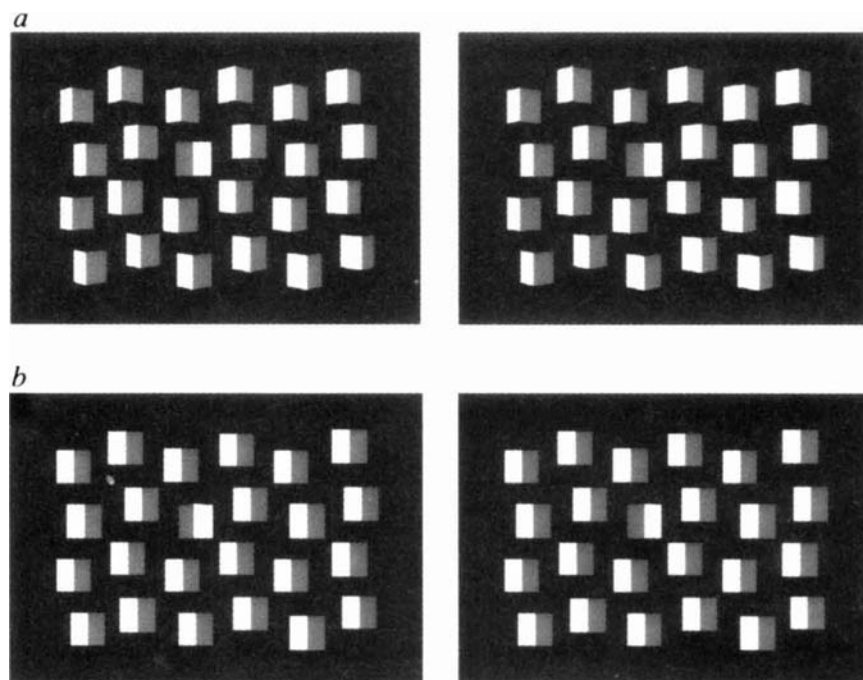
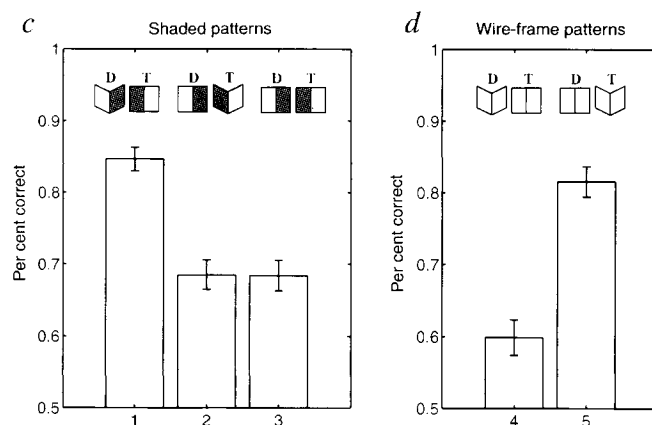


FIG. 2 Cross-fusing stereo pairs for condition 1 (a) and condition 2 (b) of experiment 1. 24 items were displayed on randomized depth planes for 180 ms, followed by a 30-ms blank screen and a 200-ms mask. One target was present among 23 distractors in 50% of the displays. Subjects reported target presence or absence at the end of each trial. c, Performance calculated from d' measurements¹⁴ averaged over 5 naive subjects for conditions 1, 2 and 3 (both distractors (D) and target (T) were flat) indicates a marked asymmetry between conditions 1 and 2. d, Wire-frame patterns were used as controls. The direction of asymmetry is opposite to that of the shaded conditions. Display duration was 250 ms, followed by masking.



different luminance patterns, the target and distractor patterns will be perceived as 3-D shapes of comparable reflectances. Hence, they will be less distinguishable from each other. As predicted, we find that the addition of 3-D shape compromises discrimination significantly (Fig. 3b).

To control for the possibility that the impairment in performance is due to spurious factors, we constructed conditions 3 and 4 (Fig. 3a). Condition 3 controls for the decrease in uniformity caused by the perspective projection; condition 4 controls for the possibility that the mere addition of 3-D shape is hampering performance. We find that neither condition 3 nor condition 4 is any more difficult than condition 1. Impairment is seen, as in condition 1, only when the luminance pattern is consistent with the 3-D shape, and may therefore be perceived as shading, rather than as changes in reflectance.

Not only does our hypothesis explain the results of experiments 1 and 2, it also explains the easy discrimination seen in Fig. 1a. In this case, we propose that the upside-down target cube, being an unusual configuration, does not drive early 3-D shape mechanisms to the degree that the familiar-looking distractors do. As a result, the target is perceived as a multiple-reflectance patch among a background of same-reflectance cubes, and is spotted with ease (Fig. 4a). This interpretation is confirmed by observers reporting that the upside-down target cube appears to have a higher luminance contrast than the distractor cubes. Our hypothesis also predicts the following: a top-lit, upright cube as target

would be difficult to spot among a multiple-reflectance background of upside-down ones. We invite the reader to test this prediction by viewing Fig. 1a upside-down (see also Fig. 4b, left). We have found under controlled experimental conditions that this 'inverted' task is pronouncedly more difficult than the original^{6,7}. A similar asymmetry is seen (Fig. 4b, right) when flat patterns that do not have any perceptual 3-D shape are used to create displays with reflectance maps equivalent to those depicted in Fig. 4b (left).

In experiment 3, we investigated whether the large performance differences seen in experiments 1 and 2 are correlated with a qualitative change from serial to parallel processing⁸. In serial visual search tasks, the time necessary for target discrimination increases significantly when the number of distractors is increased. In parallel tasks, on the other hand, the necessary time is not significantly dependent on the number of distractors. We tested the two key stimulus pairs from each of experiments 1 and 2 on two display sizes. For the stimuli from experiment 1, we found that when the target, but not the distractors, had 3-D shape, performance was serial (Fig. 4c, left). When the distractors, but not the target, had 3-D shape, performance approached parallel behaviour. For the triptych patterns from experiment 2 (Fig. 4c, right), the flat patterns elicited parallel performance whereas the 3-D patterns did not. Overall, these results suggest a qualitative change in processing behaviour that is correlated with the results of experiments 1 and 2.

Our results suggest that early vision processes compute some

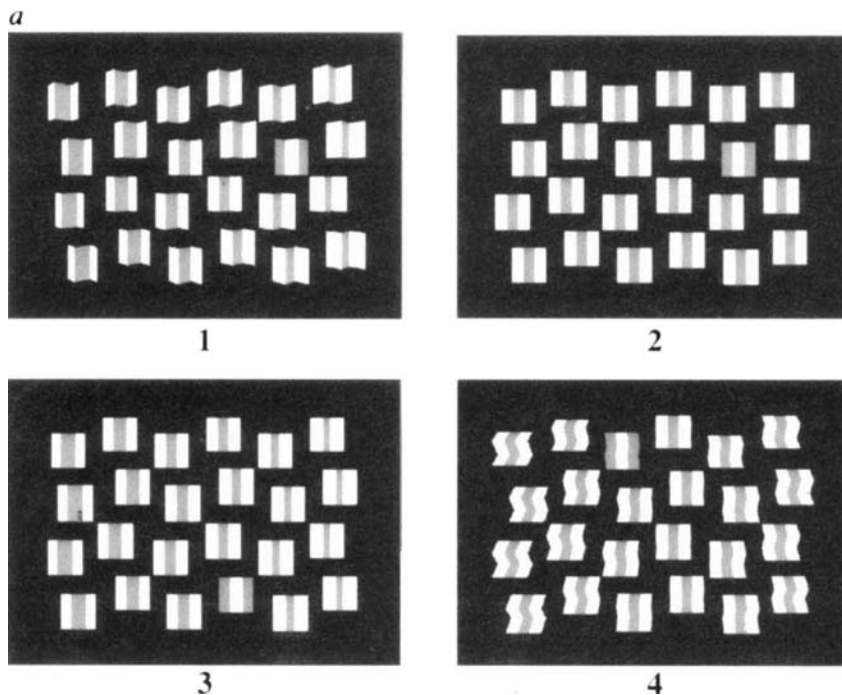


FIG. 3 *a*, The right eye's view of conditions 1–4 in experiment 2. *b*, Data were collected from 5 naive subjects performing at a display duration of 120 ms. Significant impairment of performance is seen only when the 3-D shape is perceived as single-reflectance, and the luminance pattern can be interpreted as shading (condition 1).

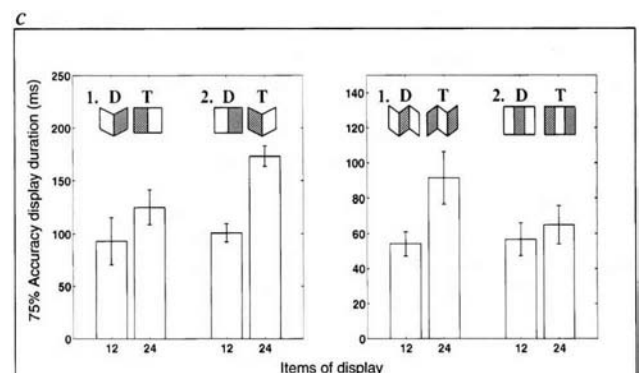
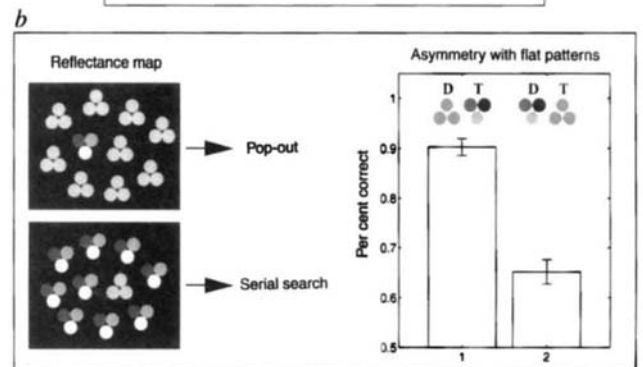
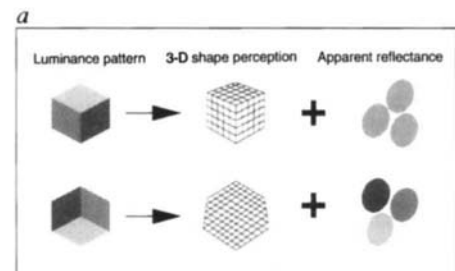
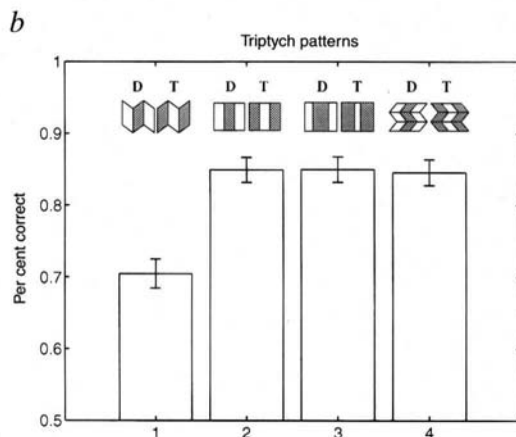


FIG. 4 *a*, The top-lit, upright cube pattern promotes 3-D shape perception by early vision processes more than the upside-down one. Such a difference in shape perception would imply a difference in reflectance perception. *b*, Left, a multiple-reflectance path would pop out from a background of single reflectance, whereas a single-reflectance patch would be difficult to spot in a multiple-reflectance background. *b*, Right, stimulus displays that have reflectance maps equivalent to those on the left, but no perceptual 3-D shape, result in a similar asymmetry in performance under controlled experimental conditions. Data were taken from 2 naive subjects at a display duration of 120 ms. *c*, 3 naive subjects were tested with displays containing either 12 or 24 stimulus patterns, using the same paradigm as in experiments 1 and 2. The display duration was varied systematically to determine the time necessary to reach 75% correct performance (data analysed as in ref. 6). Left, the stimulus patterns used were identical to those in conditions 1 and 2 in experiment 1. When the display size was increased from 12 to 24 items, there was a significant increase in the 75% accuracy display duration for condition 2 but not for condition 1. Right, stimulus patterns from conditions 1 and 2 of experiment 2 were tested in two display sizes. When the display size was increased from 12 to 24, the 75% accuracy display duration increased significantly for condition 1 but not for condition 2.

aspects of apparent reflectance together with 3-D shape. Fast and parallel discriminations are better explained on the basis of apparent reflectance than that of perceptual 3-D shape or brightness. Perceptions of 3-D shape and reflectance are closely related⁹⁻¹³; our experimental results show that this relationship begins during the very early stages of visual processing. □

Received 27 July; accepted 13 October 1995.

1. Braun, J. *Perception* **19**, A112 (1990).
2. Braun, J. *Spatial Vision* **7**, 311-322 (1993).
3. Enns, J. T. & Rensink, R. A. *Science* **247**, 71-723 (1990).
4. Enns, J. T. & Rensink, R. A. *Psychol. Rev.* **98**, 335-351 (1991).
5. Kleffner, D. A. & Ramachandran, V. S. *Percept. Psychophys.* **52**, 18-36 (1992).
6. Sun, J. Y. & Perona, P. *Invest. Ophthalmol. Vis. Sci.* **34**, 1083 (1993).
7. Sun, J. Y. & Perona, P. *Vision Res.* (in the press).
8. Treisman, A. & Gelade, G. *Cognitive Psychol.* **12**, 97-136 (1980).
9. Bergstrom, S. S. *Scand. J. Psychol.* **18**, 180-186 (1977).
10. Gilchrist, A. L. *Scient. Am.* **240**, 112-124 (1979).
11. Knill, D. C. & Kersten, D. *Nature* **351**, 228-230 (1991).
12. Metelli, F. *Scient. Am.* **230**, 91-98 (1974).
13. Adelson, E. H. *Science* **262**, 2042-2044 (1993).
14. McNicol, D. *A Primer of Signal Detection Theory* (Allen & Unwin, London, 1972).
15. Ramachandran, V. S. *Nature* **331**, 163-166 (1988).

ACKNOWLEDGEMENTS. We thank J. Braun and J. Malik for suggestions and B. Julesz, I. Kovacs, T. Sejnowski, S. Shimono and T. Watanabe for their contributions. This research was sponsored by the California Institute of Technology, an NSF National Young Investigator Award to P.P., an NSF Fellowship and Training Grant to J.Y.S., and the NSF Center for Neuromorphic Systems Engineering at Caltech.

Ataxia and epileptic seizures in mice lacking type 1 inositol 1,4,5-trisphosphate receptor

M. Matsumoto^{*†}, T. Nakagawa^{*†}, T. Inoue^{*}, E. Nagata[‡], K. Tanaka[‡], H. Takano[‡], O. Minowa[‡], J. Kuno[‡], S. Sakakibara^{*§}, M. Yamada^{*}, H. Yoneshima^{*}, A. Miyawaki^{*}, Y. Fukuchi[‡], T. Furuichi^{*}, H. Okano^{*§}, K. Mikoshiba^{*||¶}, & T. Noda[‡]

^{*} Department of Molecular Neurobiology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108, Japan

[†] Department of Cell Biology, Cancer Institute, Toshima-ku, Tokyo 170, Japan

[‡] Department of Neurology, School of Medicine, Keio University, Shinjuku-ku, Tokyo 160, Japan

[§] Molecular Neurobiology Laboratory, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), Koyadai, Tsukuba-shi, Ibaraki 305, Japan

^{||} Calciocinical Net Project, Exploratory Research for Advanced Technology (ERATO), JRDC 2-9-3 Shimo-meguro, Meguro-ku, Tokyo 153, Japan

THE inositol 1,4,5-trisphosphate (InsP₃) receptor acts as an InsP₃-gated Ca²⁺ release channel in a variety of cell types^{1,2}. Type 1 InsP₃ receptor (IP₃R1) is the major neuronal member of the IP₃R family in the central nervous system^{3,4}, predominantly enriched in cerebellar Purkinje cells but also concentrated in neurons in the hippocampal CA1 region, caudate-putamen, and cerebral cortex^{5,6}. Here we report that most IP₃R1-deficient mice generated by gene targeting die *in utero*, and born animals have severe ataxia and tonic or tonic-clonic seizures and die by the weaning period. An electroencephalogram showed that they suffer from epilepsy, indicating that IP₃R1 is essential for proper brain function. However, observation by light microscope of the haematoxylin-eosin staining of the brain and peripheral tissues of IP₃R1-deficient mice showed no abnormality, and the unique electrophysiological properties of the cerebellar Purkinje cells of IP₃R1-deficient mice were not severely impaired.

[§] Present address: Department of Molecular Neurobiology, Institute of Basic Medical Sciences, and Center of TARA, University of Tsukuba, Tsukuba, Ibaraki 305, Japan.

[¶] To whom correspondence should be addressed.

IP₃R1 gene targeting and generation of homozygous (-/-) mutant (IP₃R1^{-/-}) mice were performed essentially as described previously⁷ (Fig. 1a). Of 1,331 neomycin-resistant J1 embryonic stem cell (ES) clones, one was identified as a homologous recombinant and used to generate heterozygous (+/-) mutant mice which grew normally and had no obvious defects. Most of the IP₃R1^{-/-} mice die *in utero*, and some die postnatally. At postnatal day (PND) 10, IP₃R1^{-/-} mice were found with a frequency of 5.5% (+/+ : +/- : -/- ratio, 1.10 : 2.00 : 0.18; *n* = 271). IP₃R1^{-/-} mice were confirmed using the allele-specific polymerase chain reaction (PCR) (Fig. 1b) and Southern blot analysis (Fig. 1c). Northern blotting (Fig. 1d) showed that cerebella from IP₃R1^{-/-} mice contain a hybridizable transcript that is similar in size to the wild-type IP₃R1 messenger RNA. Reverse transcription-PCR experiments followed by Southern blot analysis using exon-specific oligonucleotide probes indicated that the transcript derived from the mutated allele lacks the first coding exon (P1; 108 bases) (data not shown). However, western blot analysis of IP₃R1^{-/-} cerebella (Fig. 1e) detected no expression of IP₃R1 at the protein level.

InsP₃ binding and InsP₃-induced Ca²⁺ release (IICR) assays were performed using cerebellar membrane fractions (P2 + P3, or microsomes). P2 + P3 fractions from IP₃R1^{-/-} cerebella exhibited a significant reduction in [³H]IP₃ binding activity compared to that from IP₃R1^{+/+} (*B*_{max} values were around 2.0 and 22.8 pmol per mg protein, respectively; Fig. 2a). In addition, almost no IICR activity was observed in IP₃R1^{-/-} microsomes, whereas ATP-dependent Ca²⁺ uptake activity was normal (Fig. 2b). These data indicate a loss of function of IP₃R1 in IP₃R1^{-/-} mice.

Body mass and overall brain sizes of IP₃R1^{-/-} mice were reduced to about 50% and 40% in weight, respectively, compared to IP₃R1^{+/+} mice, but the gross anatomy of the IP₃R1^{-/-} brains appeared normal. Staining patterns of IP₃R1^{+/+} and IP₃R1^{-/-} cerebellum (Fig. 2c), and other parts of the brain (data not shown), with haematoxylin-eosin or an anti-microtubule-associated protein-2 (MAP-2) antibody were indistinguishable. This suggests that brain structure is not significantly different at a light-microscopic level in IP₃R1^{-/-} mice. Immunohistochemical analysis of brain sections for IP₃R1 showed that it was totally absent in the IP₃R1^{-/-} brain (Fig. 2c). Moreover, quantitative autoradiography using [³H]InsP₃ demonstrated that the typical InsP₃ binding activity seen in IP₃R1^{+/+} mice was reduced in all brain regions of IP₃R1^{-/-} mice (Fig. 2d), consistent with IP₃R1 being predominantly expressed in the brain. Because a higher retention of binding was observed in the hippocampus (Fig. 2d) than in other parts of the brain, we analysed the expression level of other types of IP₃Rs in isolated hippocampal tissues. IP₃R subtypes (IP₃R1, IP₃R2 and IP₃R3) are expressed in different patterns throughout the body⁸. Western blot analysis of isolated IP₃R1^{-/-} hippocampal tissues showed no significant increase in the expression level of IP₃R2 and IP₃R3 (data not shown). The residual InsP₃-binding activity might be attributed to unknown types of IP₃R, for example, putative type 4 or type 5 IP₃R⁸, and other InsP₃-binding proteins⁹.

Gross anatomy and light-microscopic observations of paraffin-embedded sections stained with haematoxylin-eosin of IP₃R1^{-/-} tissue such as liver, pancreas, spleen, kidney, cardiac muscle, intestine and adrenal (medulla and cortex) showed no significant difference from those of IP₃R1^{-/-} mice (data not shown).

Because IP₃R1 is predominantly expressed in cerebellar Purkinje cells¹⁰, we performed electrophysiological analysis in Purkinje cells of IP₃R1^{+/+} and IP₃R1^{-/-} mice. A combination of high-frequency Na⁺ spikes and oscillatory Ca²⁺ spikes, a unique feature of the Purkinje cell¹¹, was observed in IP₃R1^{-/-} mice as well as IP₃R1^{+/+} mice (Fig. 3a). The persistent and tetrodotoxin-sensitive Na⁺ current (Fig. 3a)¹¹ and the inward rectifier current¹² (data not shown) were also observed in both types of mice. We did not observe a qualitative difference in passive and active membrane properties of Purkinje cells from the IP₃R1^{-/-} and IP₃R1^{+/+} mice.

The input-dependent graded response of excitatory postsynaptic

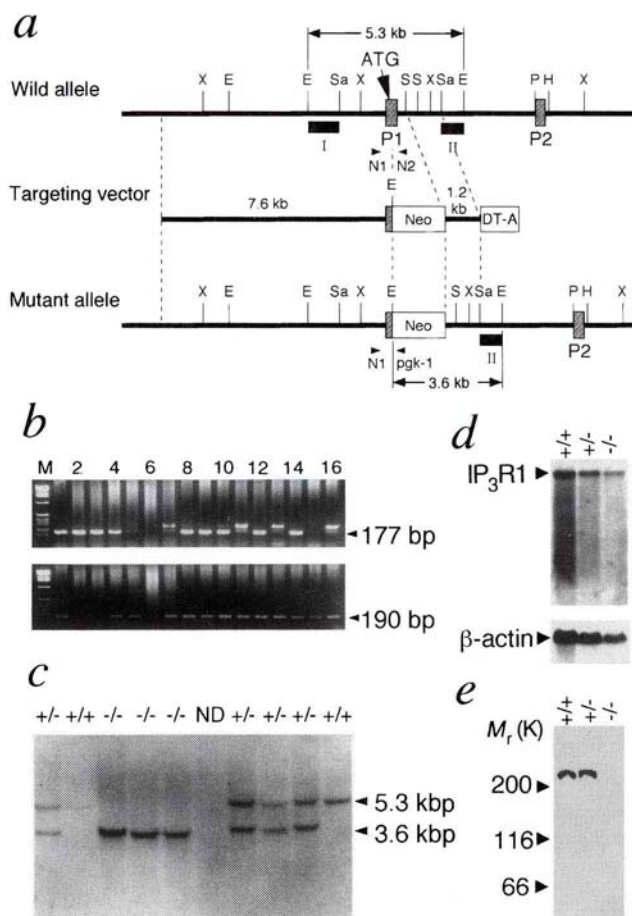
potentials (e.p.s.ps) or currents (e.p.s.cs) mediated by parallel fibres (PFs), and the all-or-none character of e.p.s.ps and e.p.s.cs mediated by climbing fibres (CFs), were observed in $IP_3R1^{-/-}$ mice as well as in $IP_3R1^{+/+}$ mice (data not shown). The time constant of PF-mediated e.p.s.c. decay, however, was significantly shorter in $IP_3R1^{-/-}$ mice (7.8 ± 2.7 ms, $n = 5$) than in $IP_3R1^{+/+}$ mice (13.2 ± 2.0 ms, $n = 6$) (mean $\pm$ s.d.; $P < 0.05$). The all-or-none character of CF-mediated e.p.s.ps suggests that $IP_3R1^{-/-}$ Purkinje cells are mono-innervated by CFs as normal Purkinje cells¹³. These excitatory synaptic transmissions were not affected by 100 μ M D(-)-2-amino-5-phosphonopentanoic acid (AP5) but were blocked by 10–50 μ M 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) in both $IP_3R1^{-/-}$ and $IP_3R1^{+/+}$ mice, indicating that the excitatory synaptic transmissions in the $IP_3R1^{-/-}$ Purkinje cells are mediated by non-NMDA (*N*-methyl-D-aspartate)-type glutamate receptors as previously described for normal rodents^{14,15}. Paired-pulse facilitation in PF-mediated e.p.s.cs and depression in CF-mediated e.p.s.cs by sequential stimuli, which are unique properties of these synapses^{15,16}, were clearly observed in all Purkinje cells from both $IP_3R1^{-/-}$ ($n = 11$ for PF, $n = 6$ for CF) and $IP_3R1^{+/+}$ mice ($n = 12$ for PF, $n = 7$ for CF) (Fig. 2b,c). The relationships between the amplitude of facilitation or depression and inter-pulse intervals were not significantly different between

$IP_3R1^{-/-}$ and $IP_3R1^{+/+}$ mice in PF-mediated e.p.s.cs (Fig. 3b), except for the 10 ms pulse interval ($P < 0.05$). However, CF-mediated e.p.s.cs in $IP_3R1^{-/-}$ mice showed a significantly stronger depression feature at pulse intervals from 20 ms to 2 s (data not shown), which could not be described quantitatively because of the incomplete space clamp (see wave forms in Fig. 3c). We conclude that the excitatory synaptic connections onto $IP_3R1^{-/-}$ Purkinje cells are not severely impaired but have some quantitative differences, which might reflect the functional roles of IP_3R1 in the Purkinje cell.

Despite the low birth rate, $IP_3R1^{-/-}$ mice appeared normal in behaviour at birth. At about PND9, they began to show ataxia, a loss of balance while standing or walking. Truncal torsions appeared at about PND 15 (Fig. 4a, left). By PND 20–23, repetitive tonic or tonic-clonic seizures prevailed, during which opisthotonus-like postures were observed (Fig. 4a, right). By intraperitoneal injections of anti-convulsants, pentobarbital (10 mg per kg, intraperitoneal) or diazepam (0.2 mg per kg, intraperitoneal), the seizures disappeared and the ataxic movements became prominent. After introduction of flumazenil (4 μ g per kg, intraperitoneal), an anti-diazepam drug, epileptic seizures suppressed by diazepam appeared again. Continuous electroencephalographic recordings of $IP_3R1^{-/-}$ mice indicated paroxysmal

FIG. 1 Targeted disruption of the gene encoding the type IP_3R . **a**, Schematic representation of the wild-type (+/+) IP_3R1 allele (top), the targeting vector (middle), and the predicted mutated allele (bottom). The first (P1) and the second (P2) coding exons are indicated by grey boxes. P1 contains a 5' untranslated sequence and the first 31 amino acid coding sequence. The targeting vector contains a PGK-neo cassette and a diphtheria toxin A (DT-A) cassette¹⁸. In this vector, a 0.7-kb fragment from the IP_3R1 gene, which includes 74 bp of P1 coding region, was deleted and replaced with a neo-resistant (neo-r) gene. Because the neo-r gene was placed just 3' to the translation initiation codon, the mutated product was expected to lack IP_3R binding activities. The final construct contains a 7.4-kb 5'-homologous region and a 1.2-kb 3'-homologous region. Arrowheads represent the positions of PCR primers; N1, 5'-TTGAGCTAAGCCAAATCTAGTGAC-3'; N2, 5'-AAGGTGCTGATAATCCATTCGTAG-3'; pgk-1, 5'-GCTAAAGCGCATGCTC-CAGACTGCCTTG-3'. Hybridization probes I and II used for screening of ES cells and mutant mice are indicated as black boxes. X, XbaI; E, EcoRI; Sa, SacI; S, SmaI; P, PstI; H, HindIII. **b**, Allele-specific PCR analysis performed with tail DNAs from offspring (F_2) of heterozygous (+/–) crosses. The 177-bp amplification product (primers N1 and pgk-1) was generated from the mutated allele (top), whereas the 190-bp product (primers N1 and N2) was generated only from the wild-type allele (bottom). Animals 1, 4, 8, 9, 10, 12 and 14 were heterozygous (+/–); 5, 7, 11, 13, 15 and 16 were wild-type (+/+); and 2 and 3 were homozygous (–/–) mutants. M, size marker. **c**, Verification of genotypes by Southern blot analysis. DNAs from $IP_3R1^{+/+}$, $IP_3R1^{-/-}$ and $IP_3R1^{+/+}$ mice were digested with EcoRI and hybridized with probe II. The wild-type allele generates a 5.3-kb band, and the mutant allele a 3.6-kb band. ND, not determined. **d**, Northern blot analysis. Poly (A)⁺ RNAs from the cerebella of $IP_3R1^{+/+}$ (lane 1), $IP_3R1^{-/-}$ (lane 2) and $IP_3R1^{+/+}$ (lane 3) mice were hybridized with IP_3R1 (top) or the β -actin (bottom) gene probe. **e**, Western blot analysis of cerebellar membrane fractions with monoclonal antibody 18A10 (ref. 19).

METHODS. **a**, Genomic clones of the mouse IP_3R1 were isolated from a library of strain 129 using an oligonucleotide probe that flanks the initiation codon of mouse IP_3R1 complementary DNA²⁰, and were sequenced. The targeting vector was constructed from a 7.6-kb genomic fragment located 5' to P1, a 1.8-kb EcoRI–SmaI neo-r gene driven by the pgk-1 promoter, a 1.2-kb genomic fragment located 3' to P1, a 1.0-kb DT-A cassette, and the plasmid pBluescript II KS- (Stratagene). The linearized targeting vector was introduced into J1 ES cells by electroporation (Bio-Rad Gene Pulser). G418 selection (175μ g ml^{−1}) was applied 36 h after the transfection, and neo-resistant colonies were isolated 7–9 days after selection. The selected ES cell colonies were cloned and their EcoRI-digested genomic DNAs were analysed by Southern blotting. Homologous recombination was proved with the 0.7-kb SacI–EcoRI fragment (probe II), and further with probe I and a neo-r probe. Of 1,331 neo-resistant clones, only one was correctly targeted. To generate chimaeric mice, this clone was injected into 3.5-day-old C57BL/6 blastocysts, and the blastocysts were implanted into pseudopregnant ICR females. Six chimaeric males with high-percentage agouti coat colour (90%) were mated with C57BL/6 females to yield heterozygous mice. The genotypes of pups obtained by crosses between the heterozygous mice were determined by PCR and/or Southern blot



analysis. **b**, Allele-specific PCR was performed in two separate reactions for each animal. **c**, Southern blot analysis was done using a standard procedure. **d**, Poly (A)⁺ RNA (2 μ g) was prepared from cerebella using Micro-Fast Track mRNA isolation kit (Invitrogen), resolved through 1.0% formaldehyde gel, and transferred onto Duralon UV membrane (Stratagene). The membrane was hybridized under high-stringency conditions with a 0.5-kb mouse IP_3R1 -specific cDNA probe²⁰. The same blots were re-probed with a human β -actin cDNA fragment. **e**, Crude membrane fractions (P2 + P3) of mouse cerebellum was prepared as described previously²¹. After SDS-PAGE (5 μ g per lane), the proteins were transferred to nitrocellulose membranes (Hybond-ECL, Amersham) and analysed by the ECL detection system (Amersham).

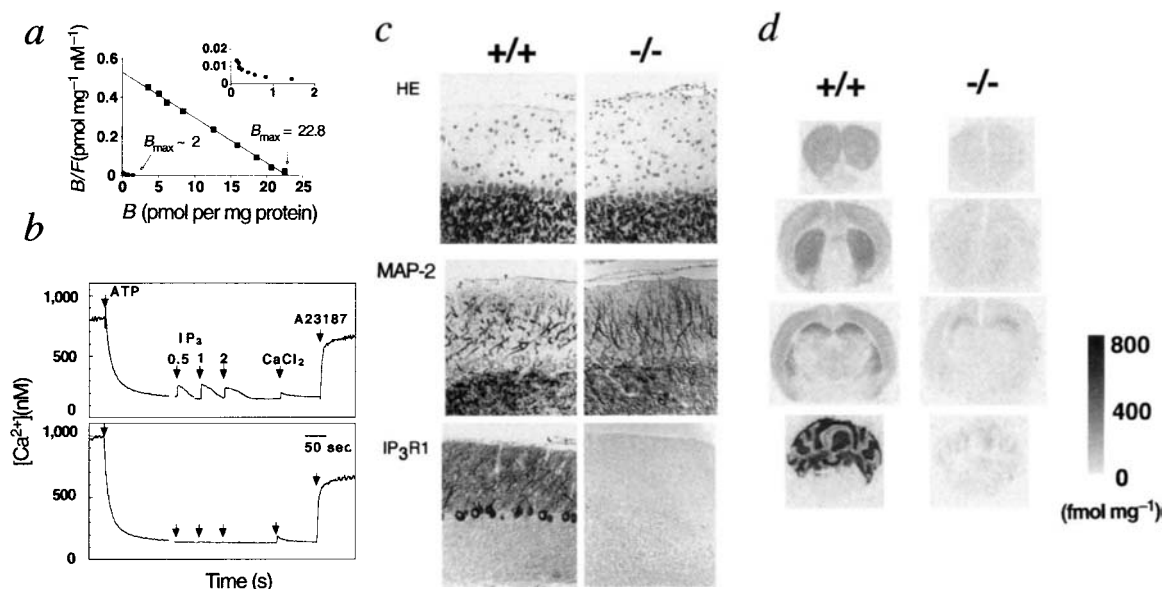


FIG. 2 Characterization of IP_3R1 deficiency in $IP_3R1^{-/-}$ mice. **a**, Saturation analysis of $[^3H]InsP_3$ binding to cerebellar membranes. Cerebellar membrane fractions (P2 + P3) from $IP_3R1^{-/-}$ (circles) and $IP_3R1^{+/+}$ (squares) littermates. Each point is a mean of duplicate determinations. B , $InsP_3$ bound to P2 + P3 fractions; F , free $InsP_3$. **b**, $InsP_3$ -induced Ca^{2+} release from cerebellar microsomes of $IP_3R1^{+/+}$ (top) and $IP_3R1^{-/-}$ (bottom) mice. Various amounts of $InsP_3$ were added at the time marked by the arrows. To confirm Fura-2 responsiveness and Ca^{2+} store functions, $CaCl_2$ (0.25 μM) and ionophore Br-A23187 (1 μM) were added at the points indicated. **c**, Histochemical analysis of $IP_3R1^{+/+}$ (left) and $IP_3R1^{-/-}$ (right) cerebella by haematoxylin–eosin stain (HE) and immunostaining with anti-microtubule-associated protein-2 (MAP-2) antibody and anti- IP_3R1 antibody (18A10 (ref. 19)). **d**, Autoradiographic localization of $InsP_3$ -binding activity. Autoradiograms of $[^3H] InsP_3$ -specific binding were obtained from $IP_3R1^{+/+}$ (left) and $IP_3R1^{-/-}$ (right) mice. A significant and diffuse reduction in the binding in the brain was noted. The $InsP_3$ binding activities of different regions of the brains of $IP_3R1^{+/+}$ and $IP_3R1^{-/-}$ mice, respectively, are shown as follows ($[^3H] InsP_3$, fmol mg^{-1} , mean $\pm$ s.d.): frontal cortex, 259 ± 42 , 83 ± 32 ;

frontoparietal cortex (motor area), 228 ± 45 , 88 ± 14 ; frontoparietal cortex (sensory area), 211 ± 30 , 77 ± 16 ; temporal cortex, 187 ± 51 , 86 ± 14 ; caudate-putamen, 357 ± 63 , 114 ± 32 ; globus pallidus, 385 ± 62 , 107 ± 32 ; thalamus (medial nuclei), 141 ± 41 , 68 ± 15 ; thalamus (lateral nuclei), 119 ± 26 , 66 ± 15 ; hippocampus CA1, 375 ± 71 , 209 ± 74 ; hippocampus CA3, 150 ± 22 , 85 ± 20 ; dentate gyrus, 181 ± 29 , 92 ± 29 ; substantia nigra, 176 ± 64 , 76 ± 24 ; cerebellar cortex, 674 ± 73 , 133 ± 29 . **METHODS.** **a**, Cerebellar membrane fractions (P2 + P3) were prepared, and $[^3H]InsP_3$ binding assays were performed as described previously^{21,22}. **b**, Preparations of cerebellar microsomes and a measurement of IICR activity using Fura-2 were as described previously²³. **c**, Animals at PND 20 were perfused with 4% paraformaldehyde. Brains were dissected and embedded in paraffin. Sagittal sections (6 μm thick) were used for immunohistochemistry. Immunohistochemistry was performed as described previously^{10,19}. **d**, Autoradiographic localization of $[^3H]InsP_3$ binding in brain sections (20 μm thick) was performed as described previously²⁴, with a slight modification, in that binding buffer of pH 8.3 and 8 nM $[^3H]InsP_3$ were used.

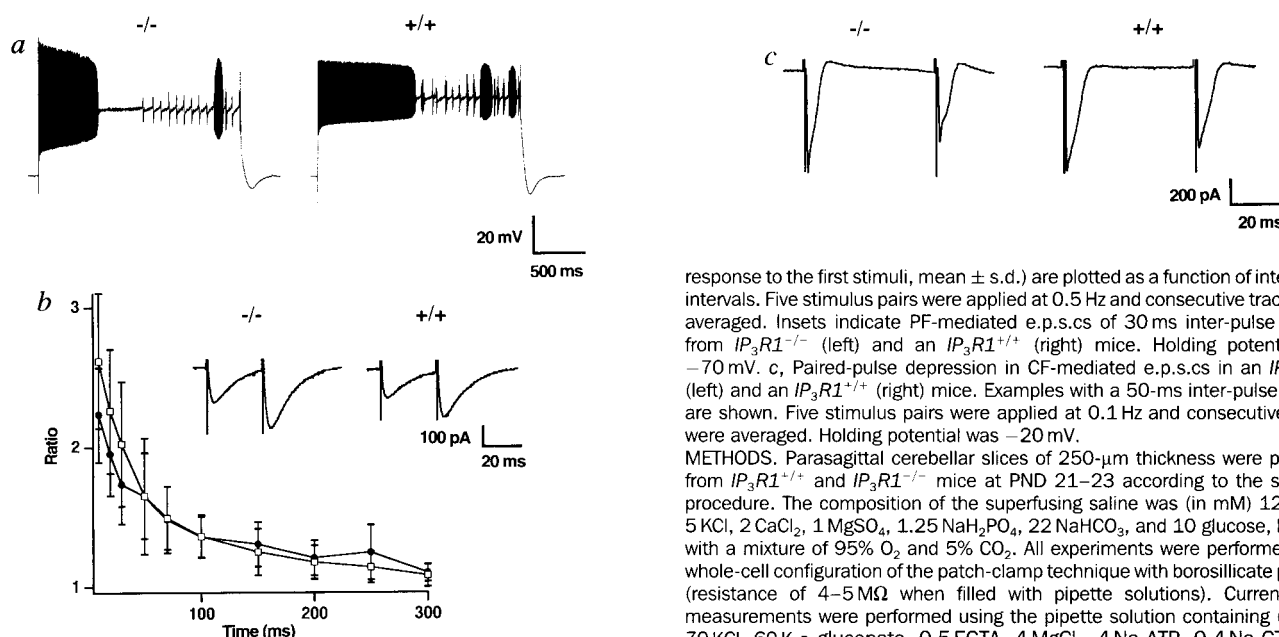


FIG. 3 Electrophysiological properties in cerebellar Purkinje cells of $IP_3R1^{-/-}$ and $IP_3R1^{+/+}$ mice. **a**, Na^+ and Ca^{2+} spikes were elicited in Purkinje cells from an $IP_3R1^{-/-}$ (left) and an $IP_3R1^{+/+}$ (right) mice by depolarizing currents injected into the somata. **b**, Paired-pulse facilitation in PF-mediated e.p.s.cs in Purkinje cells of $IP_3R1^{-/-}$ (filled circles; $n = 11$ from 5 mice) and $IP_3R1^{+/+}$ mice (open squares; $n = 12$ from 5 mice). The second responses (expressed against the

response to the first stimuli, mean $\pm$ s.d.) are plotted as a function of inter-pulse intervals. Five stimulus pairs were applied at 0.5 Hz and consecutive traces were averaged. Insets indicate PF-mediated e.p.s.cs of 30 ms inter-pulse interval from $IP_3R1^{-/-}$ (left) and an $IP_3R1^{+/+}$ (right) mice. Holding potential was -70 mV. **c**, Paired-pulse depression in CF-mediated e.p.s.cs in an $IP_3R1^{-/-}$ (left) and an $IP_3R1^{+/+}$ (right) mice. Examples with a 50-ms inter-pulse interval are shown. Five stimulus pairs were applied at 0.1 Hz and consecutive traces were averaged. Holding potential was -20 mV.

METHODS. Parasagittal cerebellar slices of 250- μm thickness were prepared from $IP_3R1^{+/+}$ and $IP_3R1^{-/-}$ mice at PND 21–23 according to the standard procedure. The composition of the superfusing saline was (in mM) 124 NaCl, 5 KCl, 2 $CaCl_2$, 1 $MgSO_4$, 1.25 NaH_2PO_4 , 22 $NaHCO_3$, and 10 glucose, bubbled with a mixture of 95% O_2 and 5% CO_2 . All experiments were performed using whole-cell configuration of the patch-clamp technique with borosilicate pipettes (resistance of 4–5 $M\Omega$ when filled with pipette solutions). Current-clamp measurements were performed using the pipette solution containing (in mM) 70 KCl, 60 K α -gluconate, 0.5 EGTA, 4 $MgCl_2$, 4 Na-ATP, 0.4 Na-GTP, and 30 HEPES, pH 7.3, with an Axoclamp 2A amplifier (Axon instruments), and voltage-clamp measurements were done using the pipette solution containing (in mM) 140 CsCl, 0.5 EGTA, 2 $MgCl_2$, 4 Na-ATP, 0.3 Na-GTP, and 10 HEPES, pH 7.3, with a CEZ-2300 amplifier (Nihon Koden). For stimulation of CFs and PFs, monopolar square pulses were applied through a glass pipette filled with the superfusing saline. Bicuculline (10 μM) was added to the saline whenever synaptically evoked events were recorded. All experiments were performed at a bath temperature of 32–34°C.

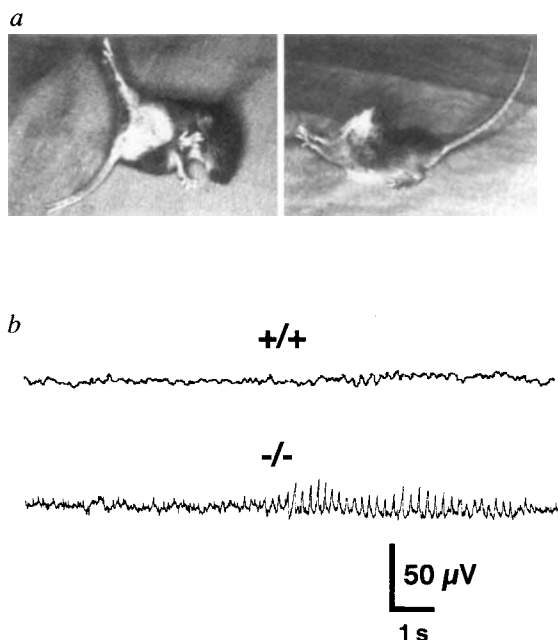


FIG. 4a, Abnormal behaviours in $IP_3R1^{-/-}$ mice at PND 20. Truncal torsion (left) and opisthotonus-like posture (right). b, Electroencephalogram of $IP_3R1^{+/+}$ and $IP_3R1^{-/-}$ littermates.

METHODS. To record electroencephalograms, mice were anaesthetized with pentobarbital (20 mg per kg, intraperitoneal). A stainless steel flat electrode was placed over the exposed parietal skull, and the reference electrode was fixed to the left ear.

polyspike activities, whereas only stable background activities were observed in $IP_3R1^{+/+}$ mice (Fig. 4b). The epilepsy and ataxia observed in $IP_3R1^{-/-}$ mice might be caused by deficits in IP_3R1 . Thus we conclude that IP_3R1 is essential for proper brain function.

It is noteworthy that the *opisthotonus* mutant locus (*Opt*, 52 cM) is located in the vicinity of the *Insp3r1* locus⁶, and that the *Opt* homozygous mutant mice exhibit phenotypes similar to those described here for $IP_3R1^{-/-}$ mice¹⁷. A complementation test of these two independent mutants should clarify this link. □

Received 25 August; accepted 27 October 1995.

- Berridge, M. J. *Nature* **361**, 315–325 (1993).
- Mikoshiba, K. *Trends pharmacol. Sci.* **14**, 86–89 (1993).
- Yamamoto-Hino, M. et al. *NeuroReport* **6**, 273–276 (1995).
- Furuichi, T., Kohda, K., Miyawaki, A. & Mikoshiba, K. *Curr. Opin. Neurobiol.* **4**, 294–303 (1994).
- Nakanishi, S., Maeda, N. & Mikoshiba, K. *J. Neurosci.* **11**, 2075–2086 (1991).
- Furuichi, T. et al. *Receptors and Channels* **1**, 11–24 (1993).
- Takeshima, H. et al. *Nature* **369**, 556–559 (1994).
- De Smedt, D. H. et al. *J. biol. Chem.* **269**, 21691–21698 (1994).
- Yagisawa, H. et al. *J. biol. Chem.* **269**, 20179–20188 (1994).
- Maeda, N., Niinobe, M., Inoue, Y. & Mikoshiba, K. *Dev. Biol.* **133**, 67–76 (1989).
- Llinas, R. & Sugimori, M. *J. Physiol., Lond.* **305**, 171–195 (1980).
- Crepel, F. & Penit-Soria, J. *J. Physiol., Lond.* **372**, 1–23 (1986).
- Eccles, J. C., Ito, M. & Szentagothai, J. *The Cerebellum as a Neuronal Machine* (Springer, Berlin, 1967).
- Llano, I., Marty, A., Armstrong, C. M. & Konnerth, A. *J. Physiol., Lond.* **434**, 183–213 (1991).
- Aiba, A. et al. *Cell* **79**, 377–388 (1994).
- Konnerth, A., Llano, I. & Armstrong, C. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2662–2665 (1990).
- Lyon, M. F. & Searle, A. G. *Genetic Variants and Strains of the Laboratory Mouse*, 2nd edn (Oxford Univ. Press, 1989).
- Yagi, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9918–9922 (1990).
- Maeda, N., Niinobe, M., Nakahira, K. & Mikoshiba, K. *J. Neurochem.* **51**, 1724–1730 (1988).
- Furuichi, T. et al. *Nature* **342**, 32–38 (1989).
- Maeda, N., Niinobe, M. & Mikoshiba, K. *EMBO J.* **9**, 61–67 (1990).
- Miyawaki, A., Furuichi, T., Maeda, N. & Mikoshiba, K. *Neuron* **5**, 11–18 (1990).
- Nakade, S., Maeda, N. & Mikoshiba, K. *Biochem. J.* **277**, 125–131 (1991).
- Nagata, E. et al. *Neuroscience* **61**, 983–990 (1994).

ACKNOWLEDGEMENTS. We thank Y. Inoue for the morphological analysis of the peripheral tissues; H. Miyakawa, T. Michikawa, J. Hirota and K. Kohda for discussions; T. Kawaguchi and D. Yasutomi for photography; and L. G. Sayers and J. M. Matheson for critical comments on the manuscript. This work was supported by the Ministry of Education, Science, Sports and Culture of Japan, and the Human Frontier Science Program.

Modulation of non-vesicular glutamate release by pH

Brian Billups & David Attwell

Department of Physiology, University College London, Gower Street, London WC1E 6BT, UK

GLUTAMATE uptake into glial cells helps to keep the brain extracellular glutamate concentration, $[glu]_o$, below levels that kill neurons. Uptake is powered^{1–4} by the transmembrane gradients of Na^+ , K^+ and pH. When the extracellular $[K^+]$ rises in brain ischaemia, uptake reverses, releasing glutamate into the extracellular space^{5,6}. Here we show, by monitoring glutamate transport electrically and detecting released glutamate with ion channels in neurons placed outside glial cells, that a raised $[H^+]$ inhibits both forward and reversed glutamate uptake. No electroneutral reversed uptake was detected, contradicting the idea⁷ that forward and reversed uptake differ fundamentally. Suppression of reversed uptake by the low pH occurring in ischaemia^{8,9} will slow the Ca^{2+} -independent release of glutamate¹⁰ which can raise $[glu]_o$ to a neurotoxic level^{11,12}, and will thus protect the brain during a transient loss of blood supply.

Glutamate uptake is partly powered by transport of OH^- ions out of the cell², so an acid shift of intracellular pH (pH_i) could inhibit uptake by depriving the carrier of internal OH^- ions. Conversely, the large acid shift of extracellular pH (pH_o) occurring in brain ischaemia^{8,9} might inhibit glutamate release by reversed uptake⁵ by depriving the carrier of external OH^- for transport into the cell¹¹. To test these ideas, electrogenic glutamate transport was studied in whole-cell clamped glial (Müller) cells isolated from the salamander retina¹³. Uptake in these cells is powered by ion movements similar^{14–18} to those in mammalian neurons and glia. Müller cells do not express significant numbers of glutamate-gated non-NMDA (*N*-methyl-D-aspartate) or NMDA receptors¹³, and in these experiments barium was present to block K^+ channels modulated by glutamatergic metabotropic receptors¹⁹. The glutamate transporters in Müller cells do not activate a significant Cl^- conductance²⁰ (unlike the cone synaptic terminal carrier and the EAAT4 carrier^{21–23}). Thus glutamate-evoked currents reflect activity of the transporters.

Applying glutamate to cells clamped with pipettes filled with solution at pH 7 generated an inward uptake current, as described previously¹³. With a more acidic pipette solution, the uptake current was smaller (Fig. 1a,b), consistent with the predicted effect of lowering the intracellular $[OH^-]$ (or with effects of pH on the uptake carrier itself). Alkalinization of the external solution inhibited the uptake current (Fig. 1c,d), consistent with a high external $[OH^-]$ preventing the loss of counter-transported² OH^- from the outer face of the carrier. The uptake current was also inhibited by external acidification, which might have been expected to increase the uptake rate by promoting the loss of OH^- from the carrier. The inhibition by low pH_o was the result of a decrease in the apparent affinity of the carrier for Na^+ (Fig. 1e), consistent with protons occupying the Na^+ -binding site and preventing transport, or altering the membrane surface potential and decreasing $[Na^+]$ near the membrane. Thus, when external and internal solutions become acidic (low pH_o and pH_i) in the first minute of brain ischaemia⁹, uptake will be inhibited.

At normal pH, when $[K^+]_o$ rises and depolarizes cells (as occurs after a few minutes of ischaemia^{9,11}), the uptake carrier is predicted to reverse and pump internal glutamate and Na^+ into the extracellular space and external K^+ and OH^- into the cell^{11,24}. This can be recorded⁵ as an outward membrane current evoked by a rise in $[K^+]_o$, which is dependent on internal Na^+ and glutamate. When the external solution was made more acidic, this K^+ -evoked current was suppressed (Fig. 2), as predicted from the lowered $[OH^-]_o$ in this situation. The reversed uptake current and its

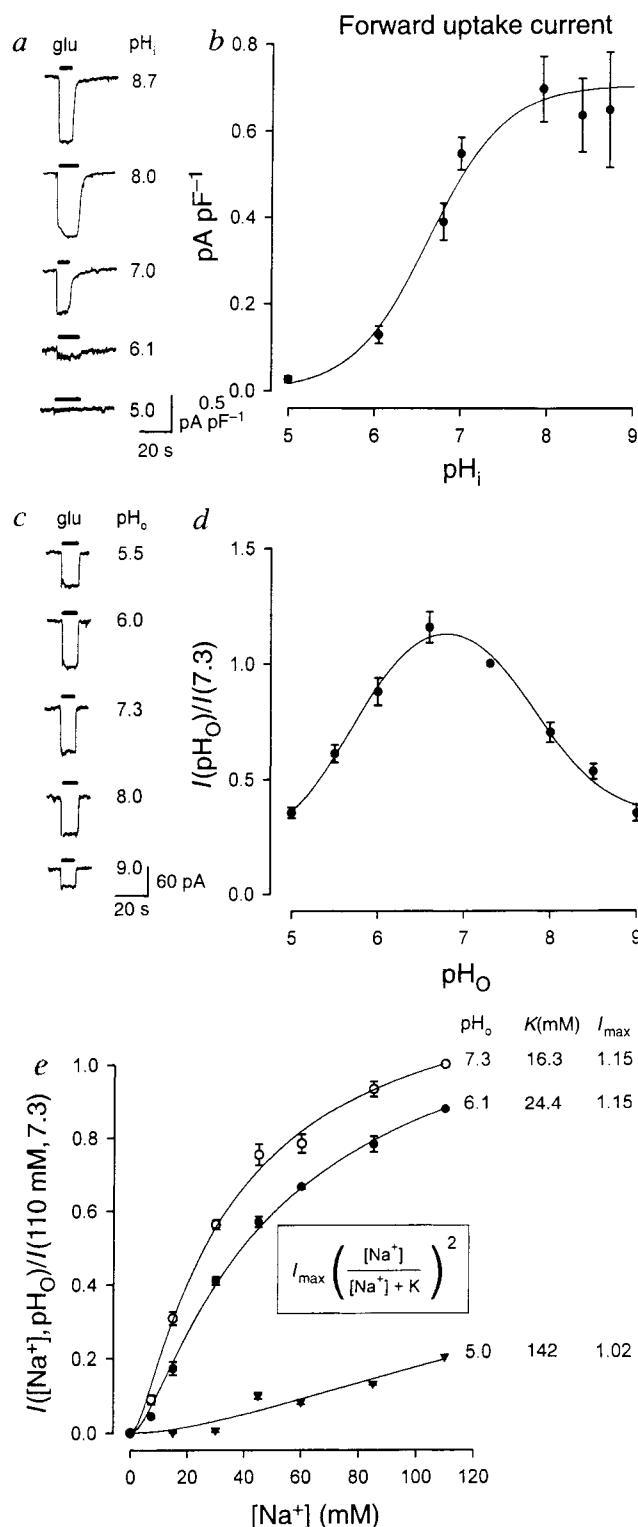


FIG. 1 Glutamate uptake is slowed by intra- and extracellular acidity. **a**, Specimen uptake currents (normalized²⁰ by cell capacitance) evoked at -40 mV by 200 μ M L-glutamate (glu) in Müller cells whole-cell clamped with solutions of different pH_i. **b**, Dependence on pH_i of the uptake current studied as in **a** (> 5 cells per point). The curve is $A[OH^-]_i/([OH^-]_i + K)$, with $K = 43$ nM and $A = 0.7$ pA pF⁻¹ (with $[OH^-] = 10^{-14}/[H^+]$). **c**, Uptake currents evoked in one Müller cell at -40 mV by 200 μ M glu in external solutions of different pH_o. **d**, pH_o-dependence of the uptake current studied as in **c** (> 5 cells per point). The curve is $1.29\{a + (1-a)/(1+10^{pH_1-pH_2})\} \{b + (1-b)/(1+10^{pH_1-pH_2})\}$, with $K_1 = 5.71$, $K_2 = 7.8$, $a = 0.14$, $b = 0.25$. **e**, [Na⁺]_o-dependence of the current evoked by 200 μ M glu at -40 mV at different pH_o (> 5 cells per point, Na⁺ replaced with choline). Fitted curves (inset) show that acidity

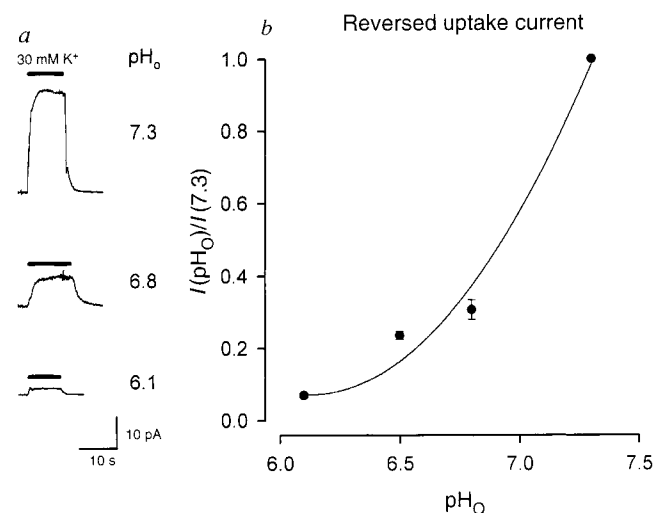


FIG. 2 External acidity suppresses the reversed glutamate uptake current. **a**, Reversed uptake currents evoked in a Müller cell at 0 mV by raising [K⁺]_o from 0 to 30 mM at different external pH (pH_o). **b**, pH_o-dependence of current in 5 cells (as in **a**, curve is a best-fit 2nd spline). To correct for small currents through unblocked K⁺ channels, the ordinate shows the K⁺-evoked current in control solution minus that in the presence of 1 mM [glu]_o. External solution (mM): KCl 30 or 0; choline-Cl 0 or 30, NaCl 75, HEPES 5, BaCl₂ 6, CaCl₂ 3, MgCl₂ 0.5, glucose 15, ouabain 0.1, pH set to 7.3 with NMDG. Internal solution (mM): Na-glu 10, choline-Cl 40, NMDG-EGTA 5, HEPES 71, NMDG 25, CaCl₂ 1, MgCl₂ 7, Na₂ATP 5, pH 7.0. With pH_i = 6.1, reversed uptake at pH_o = 7.3 was little affected (current/capacitance 0.27 ± 0.02 pA pF⁻¹ in 5 cells, compared with 0.21 ± 0.02 pA pF⁻¹ in 5 cells with pH_i = 7.0), and a similar suppression of current was seen on switching to pH_o = 6.1 (decreased to $7.3 \pm 1.5\%$ of its value at pH_o = 7.3 when pH_i = 7.0, and to $8.7 \pm 1.6\%$ when pH_i = 6.1). Slowing of reversed uptake by the acid pH shift occurring in the first minute⁹ of brain ischaemia will slow the rise of [glu]_o that occurs when [K⁺]_o rises after⁹ 2 min of ischaemia. Earlier work² suggests that a Müller cell with normal [glu]_o, [Na⁺]_o, and [K⁺]_o generates 20 pA of reversed uptake (when [K⁺]_o rises to 60 mM and depolarizes cells to -20 mV). If one glu⁻ moves per elementary charge flowing, and the retinal volume per Müller cell²⁹ is $100 \mu\text{m} \times 30 \mu\text{m} \times 30 \mu\text{m}$, of which 7% is extracellular³⁰, then this will raise [glu]_o to a neurotoxic level of 100 μ M about 3 s after [K⁺]_o rises (ignoring the decrease in release which will occur as [glu]_o rises and ignoring release by neurons). The 3–14-fold slowing of reversed uptake produced by a 0.5–1.0 unit shift of pH_o in ischaemia will prolong this to 9–42 s. Here we attribute effects of low pH to the carrier counter-transporting² OH⁻; they could equally be explained in terms of the carrier co-transporting³ H⁺.

decreases the apparent Na⁺ affinity with little effect on uptake at high [Na⁺]_o.

METHODS. Müller cells, isolated from the salamander retina¹³, were whole-cell clamped in external solution containing (in mM) KCl 2.5, NaCl 110, MgCl₂ 0.5, CaCl₂ 3, BaCl₂ 6, glucose 15, HEPES 5, pH set to 7.3 with *N*-methyl-D-glucamine (NMDG). When changing pH_i, an internal solution with (in mM) KCl 50, K₂EGTA 5, NaCl 5, Na₂ATP 5, CaCl₂ 1, MgCl₂ 7, HEPES 71, NMDG 25 was used for pH_i = 7.0; for $6.8 < \text{pH}_i < 7.9$ the amounts of HEPES and NMDG were altered. For pH_i > 7.9 or < 6.8, TAPS or MES was used as the buffer; at a given pH the buffer used did not significantly affect the results. When changing pH_o the internal solution above for pH_i = 7 was used, and external HEPES was replaced with MES for pH_o < 6.7 or TAPS for pH_o > 7.9. Monitoring² with BCECF showed that, for a pipette pH of 6 or 8, the intracellular pH was within 0.2 units of the pipette pH: in the figures pipette pH is plotted. Error bars are s.e.m.

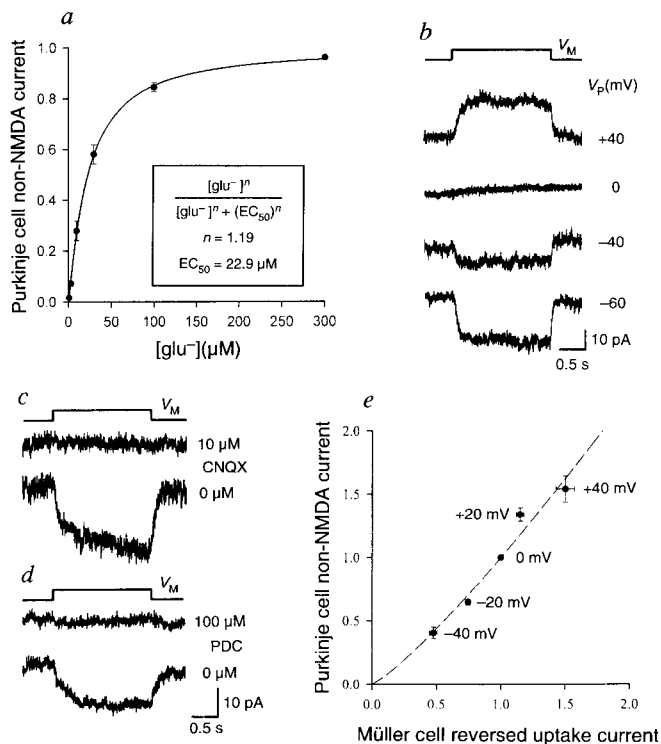


FIG. 3 Detection of glutamate release from Müller cells by reversed uptake using non-NMDA receptors in isolated Purkinje cells. **a**, Calibration of Purkinje cell $[glu^-]$ -sensitivity. Normalized currents evoked in 4 cells by different $[glu^-]$ in the presence of 1 mM trichlormethiazide (TMZ). The curve is a Hill equation (inset). **b**, Depolarizing a Müller cell containing glu^- and Na^+ from $V_M = -60$ to $+20$ mV (top) in 30 mM $[K^+]_o$ solution produced a current in an adjacent Purkinje cell which was outward at positive and inward at negative Purkinje cell potentials (V_P). **c**, **d**, The Purkinje cell current evoked at -60 mV by depolarizing a Müller cell from -60 to $+40$ mV in 30 mM $[K^+]_o$ was reversibly abolished by 10 μM CNQX (**c**, done in 4 cells) and by 100 μM PDC (**d**, 4 cells; **c** and **d** are same cells, and have the same scale bars). **e**, The Purkinje cell current evoked at -60 mV (ordinate, 8 cells) by polarizing a Müller cell in 30 mM $[K^+]_o$ from -80 mV to the potential shown by each point increased with the reversed uptake current evoked by 30 mM $[K^+]_o$ at each potential (abscissa, 8 cells). Currents are normalized to values for 0 mV. The broken curve is (reversed uptake current)^{1,19}, and shows the Purkinje cell current predicted from the calibration in **a** if glu^- release is proportional to reversed uptake current (assuming $[glu^-]$ is low micromolar as expected from the size of the reversed uptake current). **METHODS.** Purkinje cells were isolated by papain-dissociation of rat cerebellar slices. Non-NMDA receptor channels were used because they are much less pH-sensitive than NMDA receptor channels²⁸. For **a**, **b** and the ordinate of **e**, external solutions (as in Fig. 2) contained 1 mM TMZ; for **c** and **d** they contained 0.5 mM diazoxide. Pipette solution for Müller cells was as in Fig. 2, and for Purkinje cells was (mM): CsCl 110, HEPES 10, MgCl₂ 2, CaCl₂ 0.5, NMDG₂EGTA 5, Na₂ATP 2, pH set to 7.0 with NMDG.

suppression by a low pH_o were little affected when the internal pH was lowered (see Fig. 2 legend).

The assumption that the K^+ -evoked current in Fig. 2 reflects glutamate release by reversed uptake needs to be tested, because it has been claimed that reversed glutamate uptake differs fundamentally from forward uptake and is electroneutral⁷. We therefore measured glutamate release more directly, using glutamate-gated non-NMDA receptor channels in the cell bodies of isolated cerebellar Purkinje cells positioned adjacent to the Müller cells. Desensitization of the non-NMDA receptors was removed²⁵ using trichlormethiazide or diazoxide to increase their response to low $[glu]$ (Fig. 3a).

With 30 mM $[K^+]_o$, depolarizing a Müller cell containing glutamate and Na^+ (a manipulation thought to release glutamate by reversed uptake⁵) evoked a current in an adjacent whole-cell clamped Purkinje neuron. The current was inward at negative and outward at positive potentials (Fig. 3b), and was blocked by 10 μM 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) (Fig. 3c), as expected for released glutamate opening non-NMDA receptor channels. The current was blocked by the glutamate transport inhibitor L-trans-pyrrolidine-2,4-dicarboxylate (PDC) (Fig. 3d), which also produced an inward shift of Müller cell current that was larger at more positive potentials (data not shown), consistent with it blocking voltage-dependent reversed uptake⁵. Increasing

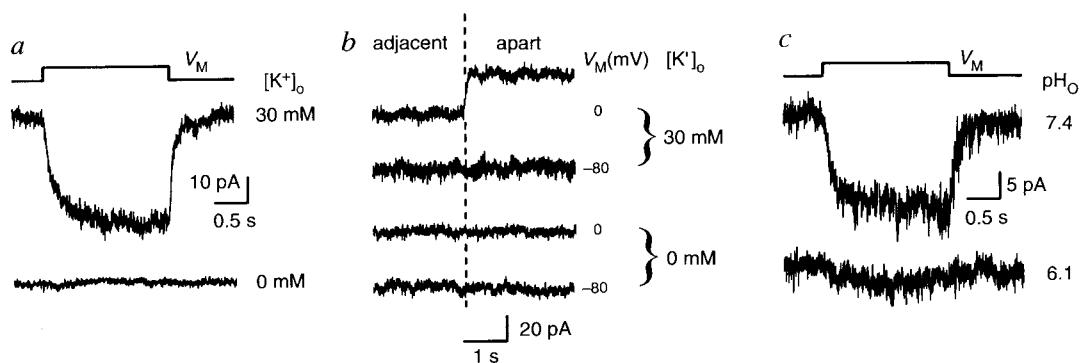


FIG. 4 Potassium-, voltage- and pH-dependence of glutamate release by reversed uptake. **a**, Depolarizing a Müller cell containing glu^- and Na^+ from $V_M = -60$ to $+20$ mV evokes a current in an adjacent Purkinje cell (at -60 mV) in 30 mM $[K^+]_o$ solution but not in zero $[K^+]_o$ (done in 9 cells). **b**, Current at -60 mV in a Purkinje cell that was initially adjacent to a Müller cell (left of traces) and was then rapidly pulled away from the Müller cell (broken line). In 30 mM $[K^+]_o$ with the Müller cell held at $V_M = 0$ mV, an upward current shift occurs when the Purkinje cell is pulled out of range of glutamate released from the Müller cell. In 30 mM $[K^+]_o$ at $V_M = -80$ mV, and in zero $[K^+]_o$ at both $V_M = 0$ and -80 mV, no current shift was seen, indicating no detectable glutamate release (done in 3 cells). **c**, The Purkinje cell current produced at -60 mV by glutamate release (evoked by Müller cell depolarization from -60 to $+20$ mV) in 30 mM $[K^+]_o$ was reduced by

$97 \pm 3\%$ at $pH_o = 6.1$ (solution buffered with MES) relative to its value at $pH_o = 7.4$ (buffered with HEPES) in 4 cells. Correcting for the lower Purkinje cell sensitivity (see below) and the dose-response curve in Fig. 3a (as in Fig. 3e), this implies a 14-fold decrease in glutamate release. A similar current reduction was seen ($87 \pm 10\%$ in 4 cells) in external solution containing 5% CO₂ and either 26 mM $[HCO_3^-]$ ($pH_o = 7.4$) or 1.2 mM $[HCO_3^-]$ ($pH_o = 6.1$). In experiments testing the sensitivity of the Purkinje cell to 3 μM glutamate, the current produced at -60 mV in zero $[K^+]_o$ was 0.94 ± 0.09 (5 cells) of that in 30 mM $[K^+]_o$, that at pH 6.1 (in 30 mM $[K^+]_o$) was 0.68 ± 0.08 without HCO_3^- present (5 cells) and 0.60 ± 0.04 with HCO_3^- present (3 cells) of that at pH 7.4, and that in the presence of PDC was 0.98 ± 0.13 (3 cells) of that in its absence. Solutions as for Fig. 3.

the Müller cell depolarization increased both the reversed uptake current in the Müller cell, as reported previously⁵, and, by a corresponding amount (according to the calibration curve in Fig. 3a), the current evoked in the Purkinje neuron by glutamate release (Fig. 3e), showing that the reversed uptake current is proportional to the rate of glutamate release.

With no external K^+ , Müller cell depolarization evoked no current change in the Purkinje neuron (Fig. 4a), consistent with reversed uptake requiring K^+ transport into the cell. To look for an electroneutral K^+ -independent component of reversed uptake⁷, which might not be evoked by depolarizing the Müller cell, we tried to detect a change in Purkinje neuron current on moving the neuron from beside the Müller cell, where it would detect glutamate released both by electrogenic and electroneutral mechanisms, to a position far away from the Müller cell. With no extracellular K^+ , no such current change was detected with the Müller cell held either at -80 mV or at 0 mV (Fig. 4b), although the same procedure with 30 mM $[K^+]_o$ produced a current change due to glutamate release with the Müller cell at 0 mV but not at -80 mV, as expected from the voltage- and $[K^+]_o$ -dependence of electrogenic reversed uptake⁵.

The Purkinje cell response to glutamate release evoked by Müller cell depolarization in 30 mM $[K^+]_o$ was reduced by changing the extracellular pH from 7.4 to 6.1 (Fig. 4c), consistent with the suppression of reversed uptake current in Fig. 2. A similar reduction was seen in the presence of external CO_2 and HCO_3^- (see Fig. 4 legend). Control experiments showed that the reduction of Purkinje cell response at pH_o 6.1 , in zero $[K^+]_o$ and in PDC, was due to a decrease in glutamate release from the Müller cell and was not due to a large reduction of Purkinje cell glutamate sensitivity (Fig. 4 legend).

Understanding reversed glutamate uptake is important because of its neurotoxic role in ischaemia. Reversed uptake has been suggested⁷ to have an electroneutral stoichiometry, in which each glutamate anion moves with one Na^+ , radically different from that of forward uptake². We found no evidence for electroneutral reversed uptake. These and previous⁵ data suggest that reversed uptake has an electrogenic stoichiometry in which $2Na^+$ and $1glu^-$ leave the cell while $1K^+$ and $1OH^-$ enter.

In stroke, $[K^+]_o$ rises to around 60 mM, leading to glutamate release by reversed uptake^{5,6,10,11}. If the oxygen deprivation is prolonged, reversed uptake is predicted²⁴ to raise $[glu]_o$ to an equilibrium level above 100 μ M, which triggers the delayed death of neurons^{11,12}. The one-unit acid shift of pH_o and pH_i occurring in ischaemia^{8,9} will not alter the predicted $[glu]_o$ reached at equilibrium, but will slow reversed uptake (Figs 2 and 4c). The existence of transient ischaemic attacks^{26,27}, when the blood supply to part of the CNS is cut off for a period of several seconds to several minutes but need not cause lasting damage, suggests that the resulting delay in the time needed for $[glu]_o$ to rise to a neurotoxic level (estimated in Fig. 2 legend) may be a useful consequence of the effect of pH on the uptake carrier. Together with the inhibition of NMDA receptor channels²⁸ by H^+ , this slowing of reversed uptake will help prevent neuronal damage during transient ischaemia. □

Received 17 August; accepted 13 November 1995.

- Kanner, B. I. & Sharon, I. *Biochemistry* **17**, 3949–3953 (1978).
- Bouvier, M., Szatkowski, M., Amato, A. & Attwell, D. *Nature* **360**, 471–474 (1992).
- Erecinska, M., Wantorsky, D. & Wilson, D. F. *J. Biol. Chem.* **258**, 9069–9077 (1983).
- Nelson, P. J., Dean, G. E., Aronson, P. S. & Rudnick, G. *Biochemistry* **22**, 5459–5463 (1983).
- Szatkowski, M., Barbour, B. & Attwell, D. *Nature* **348**, 443–446 (1990).
- Madi, J. E. & Burgess, K. J. *Neurosci.* **13**, 4429–4444 (1993).
- Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* **426**, 43–80 (1990).
- Silver, I. A. & Erecinska, M. *J. Cereb. Blood Flow Metab.* **12**, 759–772 (1992).
- Mutch, W. A. C. & Hansen, A. J. *J. Cereb. Blood Flow Metab.* **4**, 17–27 (1984).
- Ikedo, M., Nakawaza, T., Abe, K., Kaneko, T. & Yamatsu, K. *Neurosci. Lett.* **96**, 202–206 (1989).
- Szatkowski, M. & Attwell, D. *Trends Neurosci.* **17**, 359–365 (1994).
- Choi, D. W. & Rothman, S. M. *Rev. Neurosci.* **13**, 171–182 (1990).
- Brew, H. & Attwell, D. *Nature* **327**, 707–709 (1987).
- Kanai, Y. et al. *J. Biol. Chem.* **270**, 16561–16568 (1995).
- Storck, T., Schulte, S., Hofmann, K. & Stoffel, W. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10955–10959 (1992).
- Pines, G. et al. *Nature* **360**, 464–467 (1992).

- Kanai, Y. & Hediger, M. A. *Nature* **360**, 467–471 (1992).
- Attwell, D. & Mobbs, P. *Curr. Opin. Neurobiol.* **4**, 353–359 (1994).
- Schwartz, E. A. *Neuron* **10**, 1141–1149 (1993).
- Barbour, B., Brew, H. & Attwell, D. *J. Physiol., Lond.* **436**, 169–193 (1991).
- Sarantis, M., Everett, K. & Attwell, D. *Nature* **332**, 451–453 (1988).
- Eliasof, S. & Werblin, F. S. *J. Neurosci.* **13**, 402–411 (1993).
- Fairman, W. A., Vandenberg, R. J., Arriza, J. L., Kavanaugh, M. P. & Amara, S. G. *Nature* **375**, 599–603 (1995).
- Attwell, D., Barbour, B. & Szatkowski, M. *Neuron* **11**, 401–407 (1993).
- Yamada, K. A. & Tang, C.-M. *J. Neurosci.* **13**, 3904–3915 (1993).
- Nadeau, S. E. *J. Fam. Pract.* **38**, 495–504 (1994).
- Lord, R. S. *Aust. N.Z. J. Ophthalmol.* **18**, 299–305 (1990).
- Traynelis, S. F. & Cull-Candy, S. G. *Nature* **345**, 347–350 (1990).
- Mobbs, P., Brew, H. & Attwell, D. *Brain Res.* **460**, 235–245 (1988).
- Newman, E. A. & Odette, L. L. *J. Neurophysiol.* **51**, 164–182 (1984).

ACKNOWLEDGEMENTS. We thank S. Kamboj and M. Sarantis for helpful suggestions, and J. Thomas and S. Cull-Candy for providing cells. This work was supported by the Wellcome Trust, MRC and European Community.

Schwann cell apoptosis at developing neuromuscular junctions is regulated by glial growth factor

Joshua T. Trachtenberg & Wesley J. Thompson

Department of Zoology, The University of Texas, Austin, Texas 78712, USA

DENERVATED adult mammalian muscle fibres are reinnervated by regenerating axons and, in the case of partially denervated muscles, by sprouts extended from remaining, intact axons^{1–3}. Recent experiments suggest that Schwann cells (SCs) regulate these events, inducing and guiding axonal outgrowth through the processes they extend^{4,5}. In contrast to adults, reinnervation of denervated neonatal muscles is deficient and axonal sprouting is absent^{6–9}. In light of the proposed roles for SCs in these processes, we examined whether SCs in neonatal muscles exhibit altered responses to denervation. We report here that neonatal denervation leads to the rapid, apoptotic death of SCs at rat neuromuscular junctions. Injection of glial growth factor, a member of the neuregulin family of trophic factors present in developing sensory and motor neurons¹⁰, prevents this apoptosis *in vivo*. These results provide further evidence for the importance of SCs in regulating nerve growth and suggest that axon–Schwann cell trophic interactions play a role in the normal development of the neuromuscular system.

Terminal SCs, visualized with polyclonal antibodies against S100, a glial antigen, cover nerve terminals at developing neuromuscular junctions (Fig. 1a). Denervation in the early postnatal period (postnatal days 0–14 (P0–P14)), however, is followed by the rapid disappearance of terminal SCs in hind limb muscles (Fig. 1b). That SCs disappear rather than lose their expression of S100 was suggested by the lack of immunolabelling with antibodies to other Schwann cell markers, including glial fibrillary acidic protein, the low-affinity nerve growth factor receptor, and 4E2, an unidentified epitope expressed in reactive SCs¹¹. To quantify this disappearance, muscle cross-sections were double labelled with bungarotoxin to identify acetylcholine receptors at endplates and with anti-S100. Within 1 day of denervation on P4, terminal SCs could be found over $76 \pm 3\%$ of endplates (1,876 endplates, 5 muscles). By 3 days of denervation, terminal SCs remained over only $5 \pm 1\%$ of endplates (966 endplates, 3 muscles). The disappearance was not confined to SCs at the neuromuscular junction, extending to the SCs in the pre-terminal nerve branches as well. The staining intensity of the intramuscular nerves also decreased, suggesting some SCs here too were disappearing. However, the loss here was never as extensive as at the junction as many SCs persisted in these nerves.

To examine the possibility that the disappearance of SCs was a

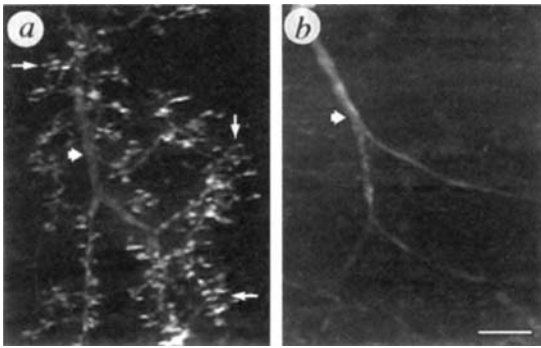


FIG. 1 Terminal SCs disappear following denervation of neonatal rat muscles. Low-magnification views of muscle whole mounts labelled with antibody to the protein S100, a marker for SCs. *a*, Innervated postnatal day 5 (P5) soleus. The SCs in the intramuscular nerve (thick arrow) and its branches, as well as the terminal SCs (thin arrows) overlying the neuromuscular junctions, are labelled. *b*, P7 soleus denervated 3 days earlier. SCs persist in the intramuscular nerve (thick arrow) but the terminal SCs have vanished. Similar results were obtained in extensor digitorum longus muscles.

METHODS. Hind limb muscles were denervated by resection of a piece of the sciatic nerve just distal to the sciatic notch in ether-anaesthetized animals. Muscles were examined following their removal from ether-killed animals by fixing with 4% buffered paraformaldehyde, permeabilizing in -20°C methanol, and labelling with a polyclonal antibody to S100 (Dako) and a TRITC-conjugated second antibody (Cappel)⁴. In some experiments, muscles were also labelled with a monoclonal antibody to neurofilament (Amersham) and a FITC-conjugated second antibody (Sigma) or with fluorochrome-conjugated bungarotoxin (Molecular Probes). Images were acquired digitally from an epi-fluorescence microscope using a Macintosh computer and an integrating CCD camera. Scale bar, 100 μm .

result of cell death, we used the TUNEL technique to identify the nuclei of cells undergoing the DNA fragmentation characteristic of apoptosis¹². In whole-mount preparations of P5 muscles denervated on P4, TUNEL-labelled nuclei in cells immunoreactive for S100 could be seen along the entire length of the intramuscular nerves and in the endplate region (Figs 2*a* and 4*a*). Apoptotic SCs were identifiable not only on the basis of their TUNEL labelling, but also by their cellular and nuclear morphology (Fig. 2*b-d*). In these muscles 1 day following denervation on P4, $27 \pm 7\%$ of remaining terminal SCs were TUNEL positive (1,289 cells, 4 muscles). Some endplates were already devoid of SCs and others had less than their normal complement, suggesting that some SCs had already completed apoptosis. Given that cells are usually identifiably apoptotic for only a short time before they are phagocytosed and disappear¹³ we believe these data provide strong evidence that apoptosis explains the disappearance of SCs.

The postnatal period during which terminal SCs die in response to muscle denervation was examined. In muscles denervated on P14 and examined on P15 $17 \pm 1\%$ (985 cells, 3 muscles) of terminal SCs were apoptotic. Three days after denervation on P14, 85% (1,605 endplates, 2 muscles) of endplates were devoid of SCs. This response was diminished in muscles denervated on P25. In these muscles $8 \pm 1\%$ (1,375 cells, 3 muscles) of terminal SCs were apoptotic after 1 day of denervation, and in muscles examined after 3 days only 5% (1,335 endplates, 2 muscles) of endplates were devoid of SCs whereas another 20% of endplates lacked their full complement of these cells. Of the remaining terminal SCs fewer than 1% showed signs of apoptosis suggesting that almost all of the apoptosis occurs within the first 3 days after denervation. It appears then that the susceptibility of SCs to denervation, rather than the rate at which these cells enter apoptosis, declines with increasing age.

The apoptosis of terminal SCs following muscle denervation suggests that, in normally developing muscle, terminal SCs are dependent upon other cell types for their survival. Two cell types

associated with terminal SCs that could provide this trophic support are motor neurons and active muscle fibres. To examine the contribution of muscle fibre activity to terminal Schwann cell survival we paralysed P4 soleus muscles with botulinum toxin ($n = 5$). In these muscles, terminal SCs persisted over endplates after 3 days of paralysis and showed no signs of apoptosis. Additionally, they extended processes (Fig. 3) that were often associated with nerve terminal sprouts (data not shown). Thus, the lack of nerve growth in response to partial denervation of neonatal muscles is likely to be due to the death of SCs rather than an inability of the axons to sprout or of SCs to induce axonal sprouting.

The absence of Schwann cell apoptosis in paralysed neonatal muscles suggests that Schwann cell viability is regulated by membrane interactions between SCs and axons, or by an axon-derived trophic factor(s) acting at short distances. Neuregulins, a family of ligands (Neu differentiation factor¹⁴, Heregulin¹⁵, glial growth factor¹⁰, ARIA¹⁶) alternatively spliced from a single gene, have been demonstrated to prevent apoptosis *in vitro* of astrocytes¹⁷ and precursors giving rise to SCs¹⁸ and are therefore attractive candidates for Schwann cell survival factors. The temporal and spatial expression of neuregulin and its putative receptor, a 185K tyrosine kinase, are consistent with such a role *in vivo*. Neuregulin messenger RNA is expressed in peripheral motor and sensory neurons beginning on embryonic day 11¹⁰ and immuno-

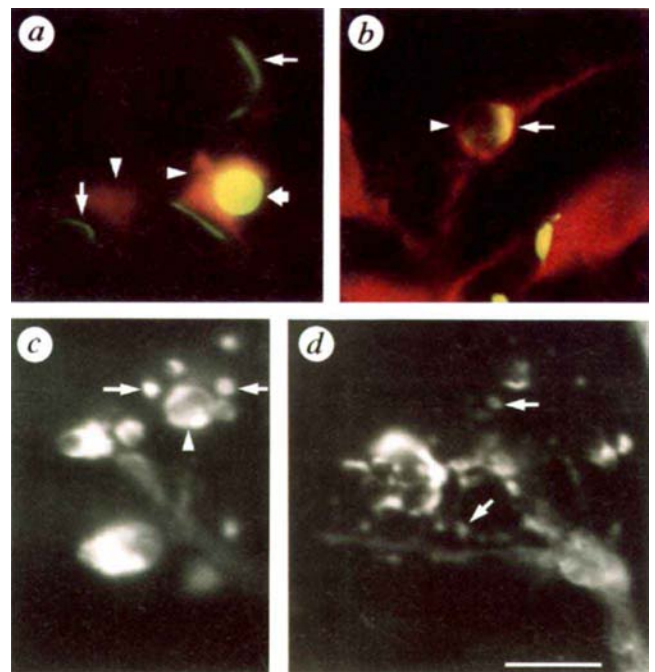


FIG. 2 Terminal SCs die by apoptosis in denervated, neonatal muscles. *a*, Cross-section of a P5 soleus muscle 1 day after denervation labelled with FITC-bungarotoxin to identify endplates (green lines at the surface of muscle fibres, thin arrows), with antibody to S100 to identify SCs (red-labelled cells at endplates, arrowheads), and with an FITC-based TUNEL technique to identify nuclei undergoing the DNA fragmentation characteristic of apoptosis (thick arrow). *b-d*, SCs show other characteristics of apoptosis 1 day after denervation. In many cases, the TUNEL-labelled chromatin was condensed at the nuclear margins (thin arrow in *b*), the S100-labelled Schwann cell cytoplasm was condensed around the nucleus (arrowhead in *b* and *c*), the Schwann cell nucleus lacked S100 labelling, and the Schwann cell was fragmenting into small S100-positive (and in some cases TUNEL-positive) pieces, presumably apoptotic bodies^{25,26} (thin arrows in *c* and *d*). In later experiments it was possible to use the S100 labelling alone to identify apoptotic cells.

METHODS. Same as in Fig. 1 except that muscles used for sections were not methanol permeabilized. The TUNEL technique used a commercial kit (Oncor, Apoptag). Scale bar, 20 μm .

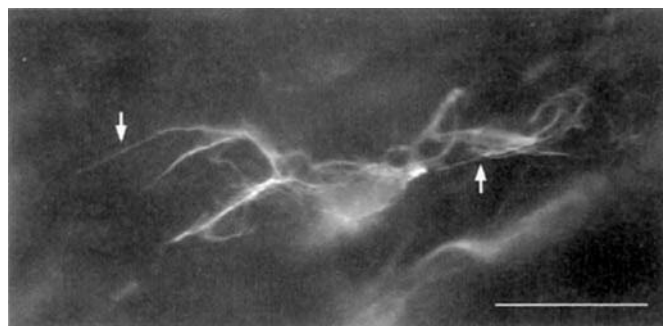


FIG. 3 Schwann cell apoptosis does not result from paralysis of neonatal muscles. An antibody to glial fibrillary acidic protein was used to label terminal SCs in neonatal muscles functionally paralysed with botulinum toxin A for 3 days. These cells persist, in contrast to their death following muscle denervation (Figs 1 and 2). Moreover, these cells also extended processes (arrows) like those extended by SCs in paralysed adult muscles⁵. METHODS. Botulinum toxin A ((2–4 ng); Calbiochem) was dissolved in 10 μ l PBS containing 0.2% gelatin. The endplate region of the soleus muscle in one hind limb was exposed in an anaesthetized, P4 rat and 10 μ l of the toxin solution applied for 10 min. The excess fluid was then removed and the wound closed with sutures. Three days later, the rat was anaesthetized and the muscle nerve stimulated *in situ*. Only muscles that failed to contract when their innervation was stimulated were processed for immunocytochemistry. Scale bar, 20 μ m.

reactivity is detectable in the motor axons of intramuscular nerves as early as P1¹⁹. Additionally, neuregulins induce tyrosine phosphorylation of the p185^{erbB2} receptor tyrosine kinase^{10,20} which is maximally expressed by SCs during the early postnatal period²¹.

To examine the possibility that neuregulins are required by developing SCs *in vivo*, recombinant human glial growth factor II (GGF), a secreted isoform of neuregulin¹⁰, was injected subcutaneously into rat hindlimbs denervated on P4. Injections of 0.4 μ g given every 8 h prevented Schwann cell apoptosis (Fig. 4a,b). In muscles denervated for 1 day, $1.0 \pm 0.5\%$ of terminal SCs were apoptotic in GGF-treated muscles (758 cells, 3 muscles) compared with $25 \pm 3\%$ in muscles exposed to vehicle alone (1,098 cells, 3 muscles). In denervated muscles exposed to GGF for 3 days (600 cells, 2 muscles), terminal SCs were present and had extended extensive processes (Fig. 4c). In contrast, denervated muscles exposed to vehicle alone were devoid of terminal SCs. Although terminal SCs persisted for 3 days after denervation in GGF-treated muscles, they were not always directly over denervated endplates, suggesting that in normally innervated muscles axons may stabilize the position of these cells.

We have presented evidence that Schwann cell survival at the developing neuromuscular junction is regulated by neurons and that this dependence extends into the fourth postnatal week. Additionally, we have identified GGF as a potential regulator of SC apoptosis *in vivo*. Previous reports have demonstrated that neuregulins can prevent the apoptosis of precursor SCs¹⁸ *in vitro*, but suggest that 'determined' SCs survive in the absence of trophic support^{18,22}. Our results demonstrate that *in vivo*, determined SCs remain trophically dependent upon neurons throughout the early postnatal period, during which time they gradually lose this dependence. This trophic dependence may play a role in the establishment and development of the neuromuscular system.

Our results also have implications for reinnervation and sprouting at developing neuromuscular junctions. In neonatal rats, some motor neurons die following temporary disconnection from their muscles²³. However, motor neurons that do not die and succeed in regenerating to their muscle nevertheless are deficient in reinnervation of muscle fibres⁹. In addition, partial denervation of neonatal muscles does not induce axonal sprouting^{7,8} as seen in the adult^{1,2}. This sprouting deficiency is unlikely to be due to the expanded size of the terminal arbors of neonatal motor neurons⁷: nerve terminals in neonatal muscle are capable of sprouting in

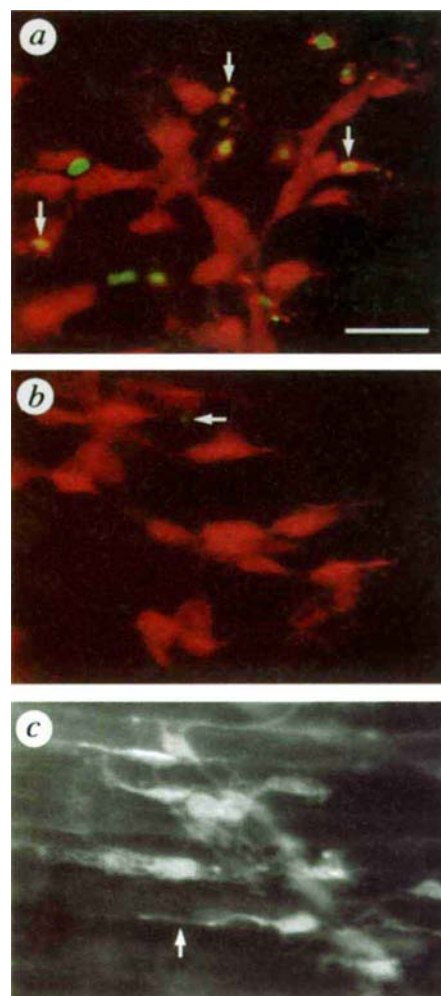


FIG. 4 Glial growth factor (GGF) rescues terminal SCs from denervation-induced apoptosis. Whole mounts of soleus muscles denervated at P4 and, at the time of the acute experiment, labelled with anti-S100 (red in a, b; white in c) and with the TUNEL technique (green/yellow in a, b). a, P5 muscle given vehicle injections for 1 day. TUNEL-labelled dying SCs (arrows) are obvious. The majority of TUNEL-labelled nuclei are in S100-expressing cells. b, P5 muscle given GGF injections for 1 day. No SCs are TUNEL-labelled. The sole labelled nucleus (arrow) in this field is not in an S100-labelled cell; most frames contained no labelled nuclei. c, P7 muscle given GGF injections for 3 days. Terminal SCs persist at a time when, in the absence of GGF, all are missing (Fig. 1). The terminal SCs have even extended processes (one indicated by arrow).

METHODS. Following denervation as described in Fig. 1 legend, rats were given subcutaneous injections (20 μ l total volume, roughly half at the hip and half at the calf of the denervated leg) every 8 h. Vehicle was PBS containing 1 mg ml⁻¹ BSA. Recombinant human GGFII (Cambridge Neuroscience, half maximum activity in an *in vitro* Schwann cell proliferation assay: 5.6 ng ml⁻¹) was dissolved in vehicle (0.02 μ g μ l⁻¹). Scale bar, 20 μ m.

response to muscle paralysis (data presented in this paper) and, in some muscles, neonatal motor units that fail to sprout are apparently smaller than these same units become after sprouting in the adult^{8,24}, suggesting that neurons in these muscles are capable of forming larger arbors. In light of the role of SCs in the adult^{4,5} we suggest that the SC death induced by denervation explains some of the deficiencies in reinnervation and sprouting in the neonate. □

Received 13 September; accepted 6 November 1995.

1. Brown, M. C. et al. *A. Rev. Neurosci.* **4**, 17–42 (1981).
2. Wernig, A. & Herrera, A. A. *Prog. Neurobiol.* **27**, 251–291 (1986).

3. Fawcett, J. W. & Keynes, R. J. *A. Rev. Neurosci.* **13**, 43–60 (1990).
4. Son, Y. J. & Thompson, W. J. *Neuron* **14**, 125–132 (1995).
5. Son, Y. J. & Thompson, W. J. *Neuron* **14**, 133–141 (1995).
6. McArdle, J. J. & Sansone, F. M. *J. Physiol.* **271**, 567–586 (1977).
7. Thompson, W. & Jansen, J. K. S. *Neuroscience* **2**, 523–535 (1977).
8. Betz, W. J., Caldwell, J. H. & Ribchester, R. R. *J. Physiol.* **303**, 265–279 (1980).
9. Dennis, M. J. & Harris, A. J. *Dev. Biol.* **74**, 173–183 (1980).
10. Marchionni, M. A. *et al. Nature* **362**, 312–318 (1993).
11. Astrow, S. H., Son, Y.-J. & Thompson, W. J. *J. Neurobiol.* **25**, 937–952 (1994).
12. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. *J. Cell Biol.* **119**, 493–501 (1992).
13. Raff, M. C. *et al. Science* **262**, 695–700 (1993).
14. Wen, D. *et al. Cell* **69**, 559–572 (1992).
15. Holmes, W. E. *et al. Science* **256**, 1205–1210 (1992).
16. Falls, D. L. *et al. Cell* **72**, 801–815 (1993).
17. Pinkas-Kramarski, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 9387–9391 (1994).
18. Dong, Z. *et al. Neuron* **15**, 585–596 (1995).

19. Sandrock, A. W. *et al. J. Neurosci.* **15**, 6124–6136 (1995).
20. Jo, S. A. *et al. Nature* **373**, 158–161 (1995).
21. Cohen, J. A. *et al. J. Neurosci. Res.* **31**, 622–634 (1992).
22. Jessen, K. R. *et al. Neuron* **12**, 509–527 (1994).
23. Snider, W. D., Elliott, J. L. & Yan, Q. J. *Neurobiol.* **9**, 1231–1246 (1992).
24. Ribchester, R. R. *J. Physiol.* **401**, 53–75 (1988).
25. Steller, H. *Science* **267**, 1445–1449 (1995).
26. Thompson, C. B. *Science* **267**, 1456–1462 (1995).

ACKNOWLEDGEMENTS. We thank C. Kirk and M. Marchionni at Cambridge Neuroscience Inc. for their generous gift of rhGGFII and advice in its use, S. Scherer for suggesting GGF as a possible trophic agent, A. Arnold for suggesting the method for labelling apoptotic cells, S. Astrow, Y.-J. Son, G. Unguez, D. Kopp and S. Scherer for their comments on the manuscript, J. Young for assistance with the artwork and L. Sutton for his technical assistance. This work was supported by a grant from the NIH.

The reaction mechanism of the internal thioester in the human complement component C4

Alister W. Dodds, Xiang-Dong Ren*, Antony C. Willis & S. K. Alex Law†

The MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

A KEY step in the elimination of pathogens from the body is the covalent binding of complement proteins C3 and C4 to their surfaces^{1–5}. Proteolytic activation of these proteins results in a conformational change^{6,7}, and an internal thioester^{8–10} is exposed which reacts with amino or hydroxyl groups on the target surface to form amide or ester bonds, or is hydrolysed^{11–15}. We report here that the binding of the human C4A isotype involves a direct reaction between amino-nucleophiles and the thioester. A two-step mechanism is used by the C4B isotype. The histidine at position 1,106 (aspartic acid in C4A) first attacks the thioester to form an acyl-imidazole intermediate. The released thiol then acts as a base to catalyse the transfer of the acyl group to amino- and hydroxyl-nucleophiles, including water.

There is ample evidence that there are thioesters in the native forms of C3 and C4 (refs 8–10), that the conformations of the molecules change dramatically on activation^{6,7}, and that the activated molecules bind covalently through the acyl group of the thioester^{11–13}, but the chemical mechanisms of the reaction are unknown. Important clues to the reaction mechanism came from the study of the two isotypes of human C4, C4A and C4B, which show different binding to substrates containing amino and hydroxyl groups^{14–16}. Once polymorphic variation is discounted, C4A and C4B differ by only four amino acids between residues 1,101 and 1,106^{17,18}. Using site-directed mutagenesis, it was established that the key residue is at position 1,106: molecules with a His at position 1,106 have C4B-like binding properties, whereas those with an Asp, Asn or Ala are C4A-like^{19,20}. These results led to the hypothesis that the His is catalytic in the reaction with hydroxyl groups, including hydrolysis. When the His is not present, hydrolysis is not catalysed and the reaction with amino groups becomes apparently more efficient. This hypothesis was confirmed by showing that the half-life ($t_{1/2}$) of activated C4A is about 10 s and that of C4B too short to be measured (< 1 s)²⁰. The His could act as a base by abstracting a proton from water to generate the hydroxide ion, which could then attack the thioester. Alternatively, His could act as a nucleophile attacking the thioester to yield an acyl-imidazole intermediate, which is subsequently hydrolysed. We therefore studied C4 variants in which H1,106 was substituted with other residues²¹. Of particular interest are C4-H1,106Y and C4-H1,106K (Table 1).

TABLE 1 Glycine and glycerol binding of natural and engineered thioester proteins

Thioester proteins	Residues 1,101–1,106	k_2/k_0 (M ⁻¹)		Relative haemolytic activity
		Glycine	Glycerol	
C4A	PCPVL	16,300	0.6	36
C4B	LSPVIH	75	16.6	100
C4-H1,106K	LSPVIK	65	0.3	1
C4-H1,106Y	LSPVIY	850	6.5	19
C3	DAPVIH	0	23.0	NA
Rat α_1 I3	SGSLIN	5,500	1.5	NA

These results were previously reported in ref. 21, except for C3 which was reported in ref. 27. The residue numbering is that of human C4. The equivalent residues of C3 and rat α_1 I3, according to sequence alignment, are also shown. The binding to glycine was done in different glycine concentrations to give non-saturating binding. The glycine concentrations were 0.1 mM for C4A and rat α_1 I3, 2.5 mM for C4B, C4-H1,106K and C3, and 1 mM for C4-H1,106Y. The concentration of glycerol was 10 mM for all proteins. The k_2/k_0 values were calculated from the binding efficiency (BE) using the formula²⁹ $k_2/k_0 = \text{BE}/[\text{G}](1-\text{BE})$ where [G] is the concentration of glycine or glycerol in the experiment, k_0 is the first-order hydrolysis rate of the thioester and k_2 the second-order rate constant of binding to glycine or glycerol. BE is the fraction of activated thioester protein bound with either glycine or glycerol and was determined experimentally. The typical concentration of the thioester proteins in these experiments, and in those reported in Figs 1 and 2, is in the range 0.5–1 mg ml⁻¹. C4 and variants were activated with C1s at a molar ratio of $\times 10$:1 and binding was carried out at 22 °C for 30 min. C3 was activated with trypsin at a molar ratio of ~ 12.5 :1 for 15 min at 37 °C, and rat α_1 I3 was activated with human neutrophil elastase at a molar ratio of ~ 60 :1 for 30 min at 37 °C. NA, not applicable.

C4-H1,106Y is the only variant, other than C4B, that binds glycerol at notable levels^{19–21}. Tyr must therefore play a similar role to His. Because the association constant (pK_a) of His (~ 6.0) and Tyr (~ 10.0) are on different sides of neutrality, it is unlikely that both assume the same role in an acid–base reaction. Thus they are likely to act as nucleophiles. The kinetics of activation and binding of C4-H1,106Y were followed. Binding lags behind activation with a $t_{1/2}$ of about 90 s (Fig. 1a). Dithionitrobenzoic acid (DTNB) was mixed with C4-H1,106Y and C1s, and release of the thiol was followed spectrophotometrically (Fig. 1a). The appearance of the thiol of C4-H1,106Y occurs soon after cleavage and before binding to glycine. An intermediate, therefore, must have been formed during the reaction.

C4-H1,106K has low haemolytic and binding activities²¹. We activated C4-H1,106K with C1s, and [³H]iodoacetic acid was included to label the released thiol at C991. After tryptic digestion, the peptides were fractionated by high-performance liquid chromatography (HPLC). One radioactive peak was found (Fig. 1b). Sequence analysis showed that it contained two peptides, with the N-terminal residues G990 and L1,081, respectively. In addition, no residue was detected at positions 994 and 1,106. These results are consistent with the hypothesis that a stable amide bond is formed between the acyl group of the thioester at Q994 and the ϵ -amino group of K1,106.

Thus the first step of the binding reaction involves the

* Present address: Department of Vascular Biology, VB-4, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

† To whom correspondence should be addressed.

FIG. 1 Properties of C4-H1,106Y and C4-H1,106K. *a*, Time course of activation (circles), binding to glycine (squares), and the appearance of the thiol (plain line) of C4-H1,106Y. Similar binding kinetics are also observed for glycerol (data not shown). *b*, The crosslinked peptides of activated C4-H1,106K. After activation by C1s in the presence of [^3H]iodoacetic acid to label the released thiol, C4-H1,106K was digested with trypsin. The HPLC elution profile of the tryptic peptides is shown. One single radioactive peak (shaded) was detected and shown to contain the peptides $\text{G}_{990}\text{CGE(Q)TMIY}$ and $\text{L}_{1,081}\text{QETSNWLLSQQADGSFQDLSPV(K)}$ in equimolar yield. No residue was identified at positions 994 and 1,106 (residues in brackets), consistent with our hypothesis that the two peptides are crosslinked at Q994 and K1,106. A similar experiment was performed with C4B (HPLC profile not shown separately). A single radioactive tryptic peptide, eluted at the position marked by the arrow, was shown to contain the sequence $\text{G}_{990}\text{CGEETMIY}$ with an E at position 994, which is expected from a hydrolysed thioester between C991 and Q994.

METHODS. *a*, The kinetics of activation and binding were determined as previously described²⁰. The appearance of the thiol was determined in a separate experiment using the same ratio of C4-H1,106Y to C1s in the presence of 100 μM 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB). The free thiol reacts with the DTNB to give the thionitrobenzoate anion, which was followed at 412 nm spectrophotometrically. The results are incorporated with those on activation and binding. *b*, The tryptic peptides of C4-H1,106K were fractionated by reversed-phase HPLC on an Aquapore OD300(C18) column (100 $\times$ 2 mm) in 0.1% trifluoroacetic acid over a linear gradient of acetonitrile to 60% at a flow rate of 200 $\mu\text{l min}^{-1}$. Each peak was collected manually. Peptide sequences were obtained on an Applied Biosystems 473A gas-phase sequencer.

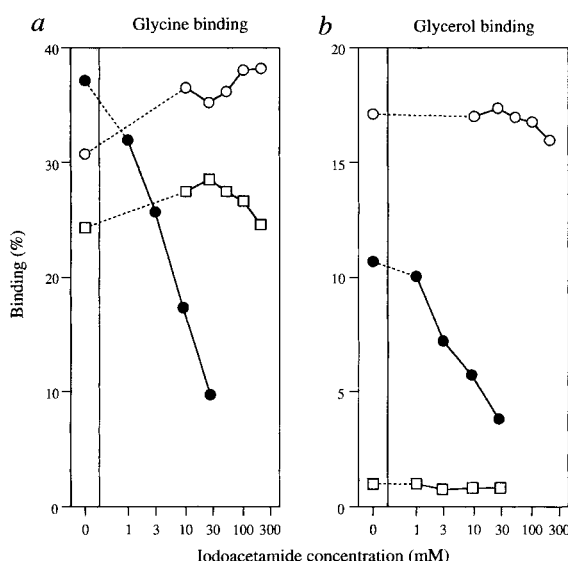
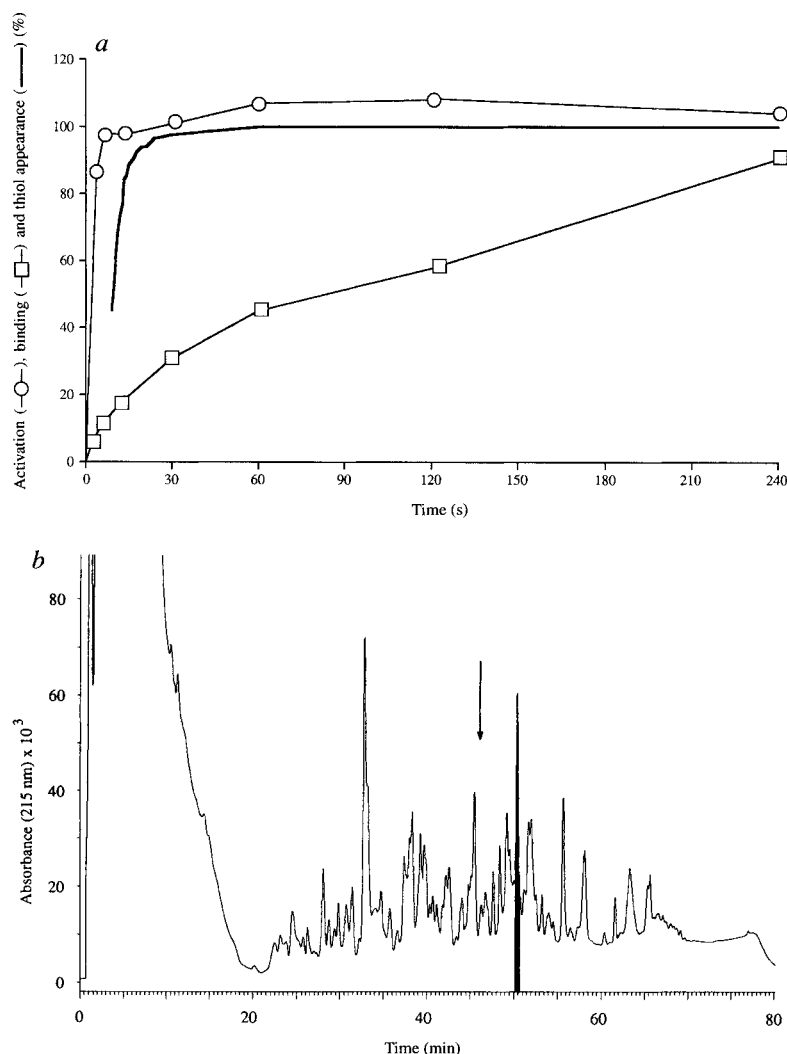


FIG. 2 The effect of iodoacetamide on the binding activity of C4A (open squares), C4B (open circles) and C4-H1,106Y (filled circles) to glycine (*a*) and glycerol (*b*). The concentrations of glycine used were 25 μM for C4A, 5 mM for C4B and 0.5 mM for C4-H1,106Y to give medium-range binding. The concentration of glycerol was 20 mM for all three proteins.

nucleophilic attack on the thioester by the residue at position 1,106. However, two problems remain: (1) a base is still required to catalyse the reaction of the acyl-intermediates with substrate molecules; and (2) the reaction seems unnecessarily complicated: why should one activated acyl group (thioester) be replaced by another (the acyl-intermediates) in the reaction? Both problems could, however, be overcome by postulating that the thiolate anion of C991, released while forming the intermediate, is the base. The formation of the intermediate becomes an obligatory step to keep the acyl group activated while releasing the thiol group. Iodoacetamide, which reacts specifically with thiolate anions, was included in the binding reaction of C4A, C4B and C4-H1,106Y with substrates. Binding of C4A and C4B were not affected but inhibition was observed for C4-H1,106Y (Fig. 2).

We now have a model (Fig. 3) for the covalent binding reaction of C4 that can account for all observable properties of the various natural C4 allotypes^{14–16} and variants constructed by ourselves^{20,21} and others¹⁹. The key residue is at position 1,106. If it is non-nucleophilic (Ala, Asn, Ser, Arg or Asp as in C4A), the thioester reacts directly with substrate nucleophiles. In these cases, the reaction with the more nucleophilic amino groups is dominant and hydrolysis is relatively slow. Thiol-specific reagents have no effect on binding because the thiol of C991 is released only when binding occurs. If the residue at position 1,106 is nucleophilic (Lys, Tyr or His as in C4B), it attacks the thioester and forms an acyl-intermediate. The released thiolate anion of C991 then acts as a base to catalyse the second part of the reaction between the

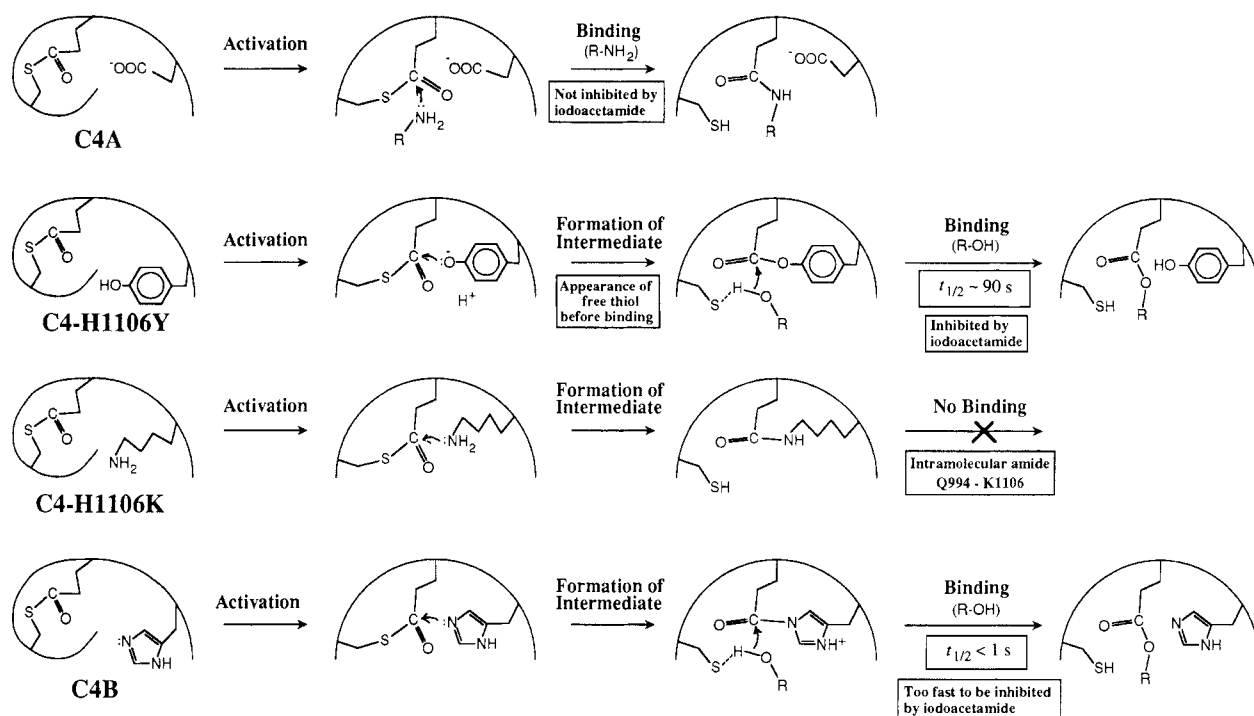


FIG. 3 The chemical mechanisms of activation and binding of C4A, C4-H1,106Y, C4-H1,106K and C4B. Key properties essential for the formulation of the reaction mechanisms are boxed.

acyl-intermediate and substrates including water. The binding reaction of C4-H1,106Y ($t_{1/2}$ of 90 s) is inhibited by iodoacetamide in the range of 1–30 mM. The binding reaction of C4B is much faster ($t_{1/2} < 1$ s), and cannot be inhibited by iodoacetamide up to 200 mM (Fig. 2). An intramolecular amide bond is formed when C4-H1,106K is activated and it is stable to further reaction. The binding of glycine observed at high substrate concentrations (Table 1) probably reflects competition between glycine and the ϵ -amino group of K1,106 for the thioester in the initial stage of the reaction.

The α -macroglobulin family of protease inhibitors also have internal thioesters²² and the residue equivalent to C4-1,106 is, in most cases, Asn. Their binding activities are complicated by their being mainly dimeric or tetrameric and it is considered that conformational changes to 'trap' proteases are more important than covalent binding as the inhibitory mechanism²². However, a monomeric α -macroglobulin from the rat, α_1 I3, whose function requires covalent binding²³ and which has an Asn at the relevant position^{24,25}, is very similar to C4A in its binding properties (Table 1)²¹. C3, which has a His at the position equivalent to C4-1,106 (ref. 26), binds preferentially to hydroxyls, and its binding to amino groups is minimal at neutral pH (Table 1)²⁷.

Thus our model may not be complete: other residues may also influence the binding reaction to account for the difference, for example, between C3 and C4B. In addition, it provides no information on the conformational change during activation, which is probably required to bring residue 1,106 close to the thioester. These points will be clarified only when we have three-dimensional structural data on the native and activated molecules.

Of the two C4 isotypes, C4B is the more universal form because, although C4B has been found in all species of mammals studied, C4A was found, in addition to C4B, only in primates, sheep and cattle^{28,29}. The covalent binding reaction of C4B, possibly also that of C3, must therefore be considered more important. The mechanism proposed ensures a rapid reaction with a built-in safety feature in the hydrolysis of the thioester, so that activated molecules cannot diffuse away from the site of activation to bind to bystander cells of the host. Finally, it has long been known that the free thiols of the activated forms of these proteins are subject to rapid oxidation and their quantitative titration requires the immediate presence of thiol-specific reagents³⁰. The thioester is therefore an ingenious device serving the dual purpose of keeping the acyl group ready for activation and protecting the thiol group from oxidation. □

Received 18 July; accepted 1 November 1995.

- Law, S. K. A. & Reid, K. B. M. *Complement: in Focus Series* 2nd edn (ed. Male, D.) (IRL, Oxford, 1995).
- Levine, R. P. & Dodds, A. W. *Curr. Topics Microbiol. Immun.* **153**, 73–82 (1989).
- Ng, Y. C. & Schifferli, J. A. in *Complement in Health and Disease* 2nd edn (eds Whaley, K., Loos, M. & Weiler, J. M.) 199–228 (Kluwer Dordrecht, Boston, 1993).
- Weitzman, J. B. & Law, S. K. A. in *Complement in Health and Disease* 2nd edn (eds Whaley, K., Loos, M. & Weiler, J. M.) 269–297 (Kluwer Dordrecht, 1993).
- Fearon, D. T. *Current Opin. Immun.* **5**, 341–348 (1993).
- Isenman, D. E., Kells, D. I. C., Cooper, N. R., Müller-Eberhard, H. J. & Pangburn, M. K. *Biochemistry* **20**, 4458–4467 (1981).
- Isenman, D. E. & Kells, D. I. C. *Biochemistry* **21**, 1109–1117 (1982).
- Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L. & Prahl, J. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5764–5768 (1980).
- Campbell, R. D., Gagnon, J. & Porter, R. R. *Biochem. J.* **199**, 359–370 (1981).
- Pangburn, M. K. & Müller-Eberhard, H. J. *J. exp. Med.* **152**, 1102–1114 (1980).
- Law, S. K. & Levine, R. P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2701–2705 (1977).
- Law, S. K., Lichtenberg, N. A. & Levine, R. P. *J. Immun.* **123**, 1388–1394 (1979).
- Hostetter, M. K., Thomas, M. L., Rosen, F. S. & Tack, B. F. *Nature* **298**, 72–75 (1982).
- Isenman, D. E. & Young, J. R. *J. Immun.* **132**, 3019–3027 (1984).

- Law, S. K. A., Dodds, A. W. & Porter, R. R. *EMBO J.* **3**, 1819–1823 (1984).
- Dodds, A. W., Law, S. K. A. & Porter, R. R. *Immunogenetics* **24**, 279–285 (1986).
- Belt, K. T., Carroll, M. C. & Porter, R. R. *Cell* **36**, 907–914 (1984).
- Yu, C. Y., Belt, K. T., Giles, C. M., Campbell, R. D. & Porter, R. R. *EMBO J.* **5**, 2873–2881 (1986).
- Carroll, M. C., Fathallah, D. M., Bergamaschini, L., Alicot, E. M. & Isenman, D. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6868–6872 (1990).
- Sepp, A. et al. *Protein Sci.* **2**, 706–716 (1993).
- Ren, X. D., Dodds, A. W., Enghild, J. J., Chu, C. T. & Law, S. K. A. *FEBS Lett.* **368**, 87–91 (1995).
- Chu, C. T. & Pizzo, S. V. *Lab. Invest.* **71**, 792–812 (1994).
- Enghild, J. J., Salvasen, G., Thøgersen, I. B. & Pizzo, S. V. *J. Biol. Chem.* **264**, 11428–11435 (1989).
- Schweizer, M. et al. *Eur. J. Biochem.* **164**, 375–381 (1987).
- Braciak, T. A. et al. *J. Biol. Chem.* **263**, 3999–4012 (1988).
- de Bruijn, M. H. L. & Fey, G. H. *Proc. Natn. Acad. Sci. U.S.A.* **82**, 708–712 (1985).
- Dodds, A. W. & Law, S. K. A. *Complement* **5**, 89–97 (1988).
- Dodds, A. W. & Law, S. K. A. *Biochem. J.* **265**, 494–502 (1990).
- Ren, X. D., Dodds, A. W. & Law, S. K. A. *Immunogenetics* **37**, 120–128 (1993).
- Thomas, M. L. & Tack, B. F. *Biochemistry* **22**, 942–947 (1983).

An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast

Julie H. Cocker*, Simonetta Piatti†, Corrado Santocanale*, Kim Nasmyth† & John F. X. Diffley*‡

* Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK

† Research Institute of Molecular Pathology, Dr. Bohr Gasse 7, 1030 Vienna, Austria

‡ To whom correspondence should be addressed

ORIGINS of DNA replication in *Saccharomyces cerevisiae* are bound by two protein complexes during the cell cycle^{1,2}. Post-replicative complexes closely resemble those generated *in vitro* by purified origin recognition complex (ORC)^{1,3-5}, which is essential for DNA replication *in vivo*⁶⁻¹¹. Pre-replicative complexes (pre-RCs) are characterized by an extended region of nuclease protection overlapping the ORC footprint¹. We show here that the Cdc6 protein (Cdc6p), which is necessary for origin firing *in vivo*¹²⁻¹⁶, is essential for the establishment and maintenance of pre-RCs, suggesting that it is a component of these complexes. Without Cdc6p, G1 origins closely resemble post-replicative origins, providing evidence that ORC is also a component of pre-RCs. These results suggest that pre-RCs play an essential role in initiating DNA replication and support a two-step mechanism for the assembly of functional initiation complexes.

Cdc6p has been implicated in an early event in S phase^{9,12-17} and is expressed in actively cycling cells in late mitosis at roughly the time of pre-RC formation^{15,18}. To examine the role of Cdc6p in pre-RC formation, we used a strain with the *CDC6* gene under control of the regulatable *MET3* promoter¹⁵ (Fig. 1a). Cells were first uniformly blocked in late anaphase using a temperature-sensitive *cdc15* mutation; half the culture was subsequently transferred to methionine-containing medium (+MET) which represses the *MET3* promoter¹⁹ and then released into G1 (+MET) while the other half was transferred to fresh methionine-free (-MET) medium and released into G1 (-MET). The state of the 2- μ m origin was examined by genomic footprinting at the indicated times. During the *cdc15* block, replication origins are in the post-replicative state (Fig. 1b, lanes 1-3)

characterized by protection of the essential ARS consensus sequence (ACS) and the presence of an ORC-induced hypersensitive site¹ (Fig. 1b, asterisk). Under conditions where *CDC6* is expressed (-MET), pre-RCs, characterized by extended nuclease protection (region of greatest protection shown by bracket, Fig. 2b) including suppression of the ORC-induced DNase1 hypersensitive site, formed between 30 and 60 min after release (Fig. 1b, lanes 4-9). Under conditions where *CDC6* expression is repressed (+MET), however, pre-RCs never formed as evidenced by the persistence of the ORC-induced hypersensitive site and lack of protection of the flanking sequence (Fig. 1b, lanes 10-15).

The experiment in Fig. 2 demonstrates that release from mitosis into methionine-containing medium blocks pre-RC formation in the *MET3-CDC6* strain (Figs 1 and 2b, lanes 7-12) but not in a strain in which *CDC6* is expressed from its natural promoter (Fig. 2b, lanes 1-6). Thus, the inability of the *MET3-CDC6*-containing strain to form pre-RCs in the presence of methionine is a consequence of reduced *CDC6* expression and not due merely to the presence of methionine. Together, these experiments demonstrate that *de novo* Cdc6p synthesis is required for the

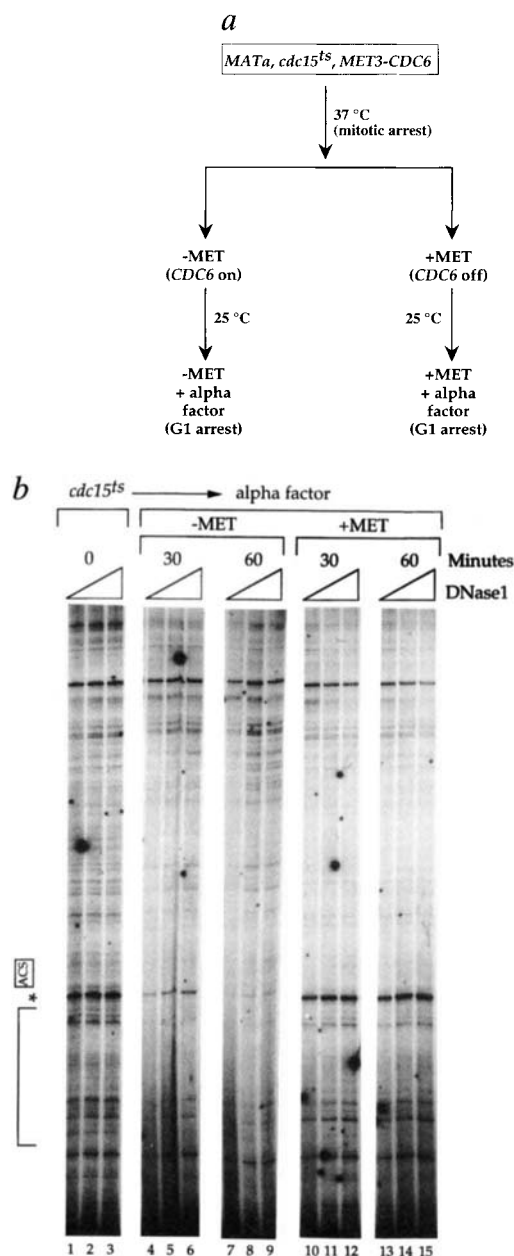


FIG. 1 A *MET3-CDC6* strain cannot form pre-RCs in the presence of methionine. The experiment in *b* is outlined in *a*. In this experiment, a culture of 4142 cells (*MATa cdc15-2, cdc6::hisGURA3hisG, trp::TRP1 MET-CDC6, leu2*)¹⁵ was first arrested in late anaphase at the *cdc15* non-permissive temperature (37 °C). Half of the culture was then transferred to synthetic complete medium without methionine (-MET) at 37 °C while the other half was transferred to synthetic complete medium with 2 mM methionine (+MET) at 37 °C. After 30 min, each culture was transferred to fresh synthetic complete medium containing alpha factor either with or without methionine at the *cdc15* permissive temperature (25 °C). At the indicated times, cells were collected and processed for genomic footprinting. For each time point, a titration of increasing DNase1 amounts is shown, indicated as a triangle. The position of the ORC-induced hypersensitive site is indicated as an asterisk whereas the most prominent additional protection generated by the pre-RC is indicated by the bracket.

METHODS. Cells were pre-grown in synthetic complete medium lacking methionine²⁶ to a density of $\sim 1 \times 10^7$ cells per ml and transferred to fresh medium pre-warmed to 37 °C. Synthetic alpha factor was used at $10 \mu\text{g ml}^{-1}$. All genomic footprinting and primer extension reactions were performed essentially as previously described using the 2- μ m-specific primer 60¹.

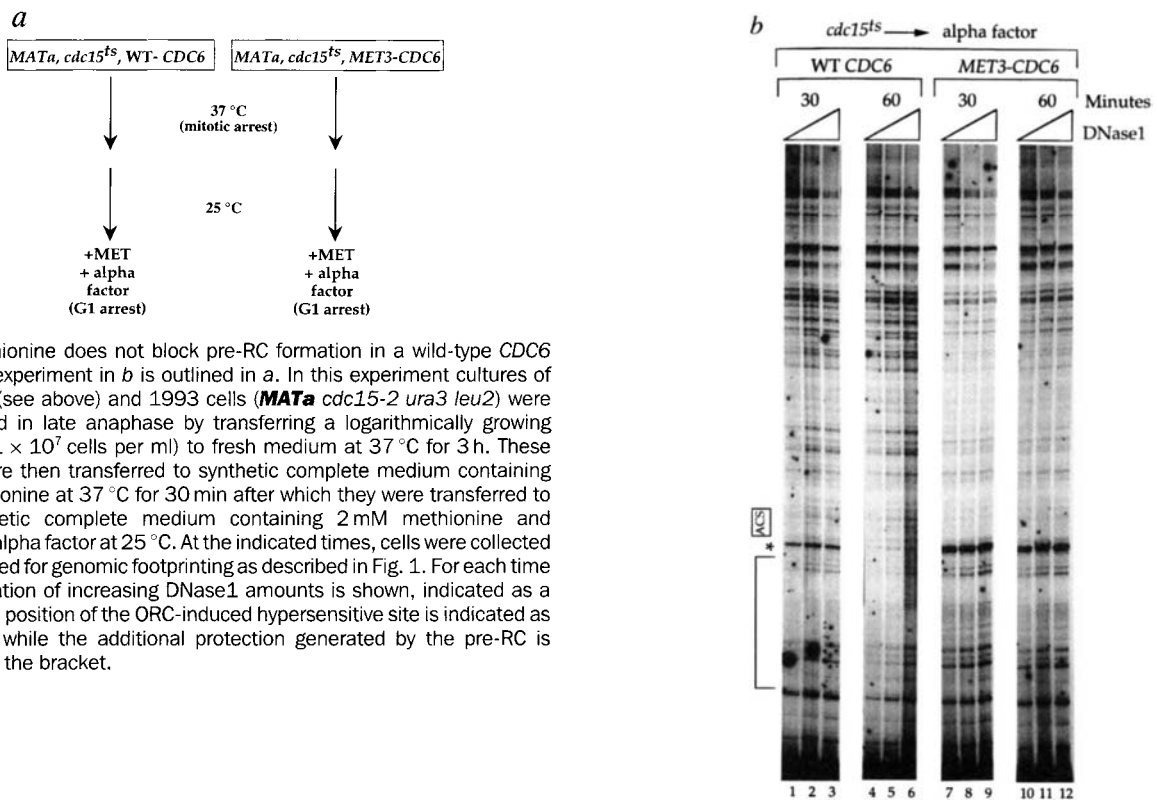


FIG. 2 Methionine does not block pre-RC formation in a wild-type *CDC6* strain. The experiment in *b* is outlined in *a*. In this experiment cultures of 4142 cells (see above) and 1993 cells (*MATa cdc15-2 ura3 leu2*) were first arrested in late anaphase by transferring a logarithmically growing culture ($\sim 1 \times 10^7$ cells per ml) to fresh medium at 37 °C for 3 h. These cultures were then transferred to synthetic complete medium containing 2 mM methionine at 37 °C for 30 min after which they were transferred to fresh synthetic complete medium containing 2 mM methionine and 10 $\mu\text{g ml}^{-1}$ alpha factor at 25 °C. At the indicated times, cells were collected and processed for genomic footprinting as described in Fig. 1. For each time point, a titration of increasing DNaseI amounts is shown, indicated as a triangle. The position of the ORC-induced hypersensitive site is indicated as an asterisk while the additional protection generated by the pre-RC is indicated by the bracket.

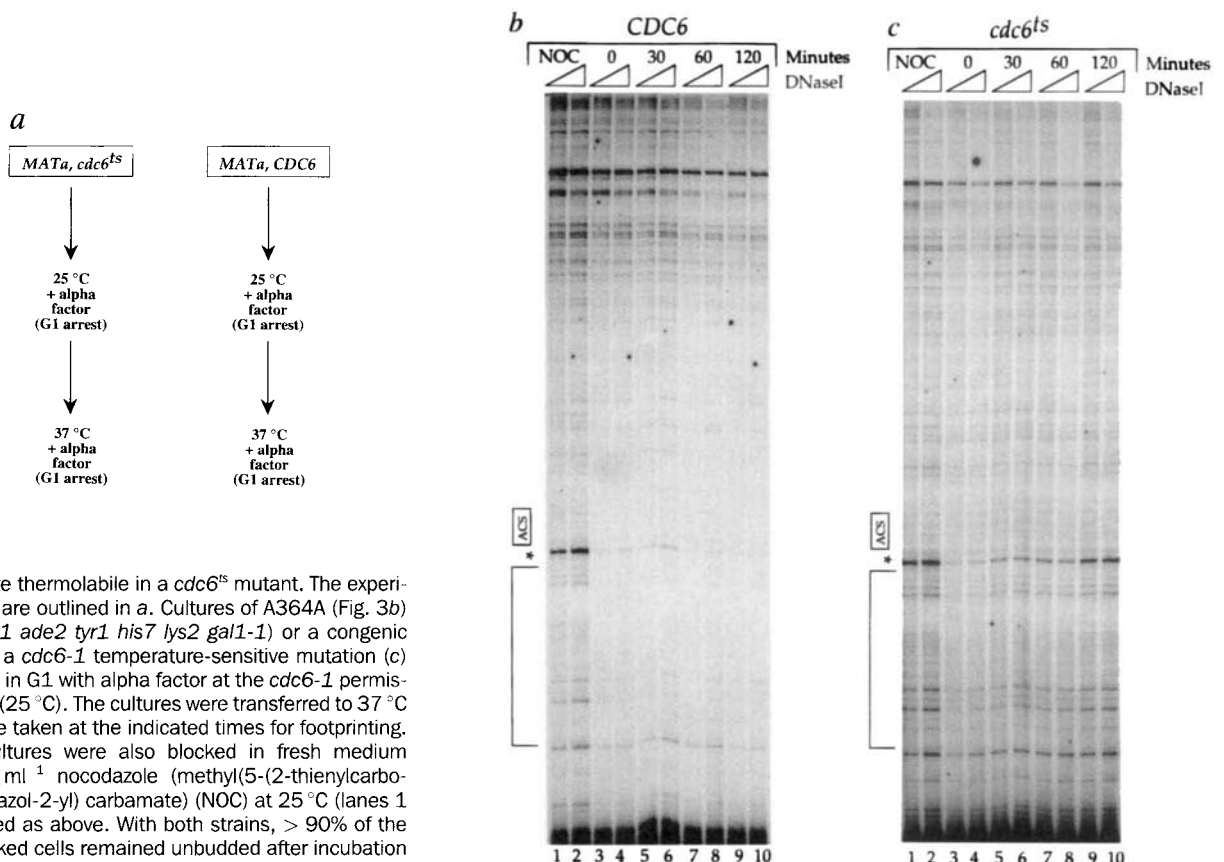


FIG. 3 Pre-RCs are thermolabile in a *cdc6^{ts}* mutant. The experiments in *b* and *c* are outlined in *a*. Cultures of A364A (Fig. 3b) (*MATa ura1 ade1 ade2 tyr1 his7 lys2 gal1-1*) or a congenic strain harbouring a *cdc6-1* temperature-sensitive mutation (*c*) were first blocked in G1 with alpha factor at the *cdc6-1* permissive temperature (25 °C). The cultures were transferred to 37 °C and samples were taken at the indicated times for footprinting. As a control, cultures were also blocked in fresh medium containing 20 $\mu\text{g ml}^{-1}$ nocodazole (methyl(5-(2-thienylcarbonyl)-1H-benzimidazol-2-yl) carbamate) (NOC) at 25 °C (lanes 1 and 2) and treated as above. With both strains, > 90% of the alpha-factor-blocked cells remained unbudded after incubation at the non-permissive temperature.

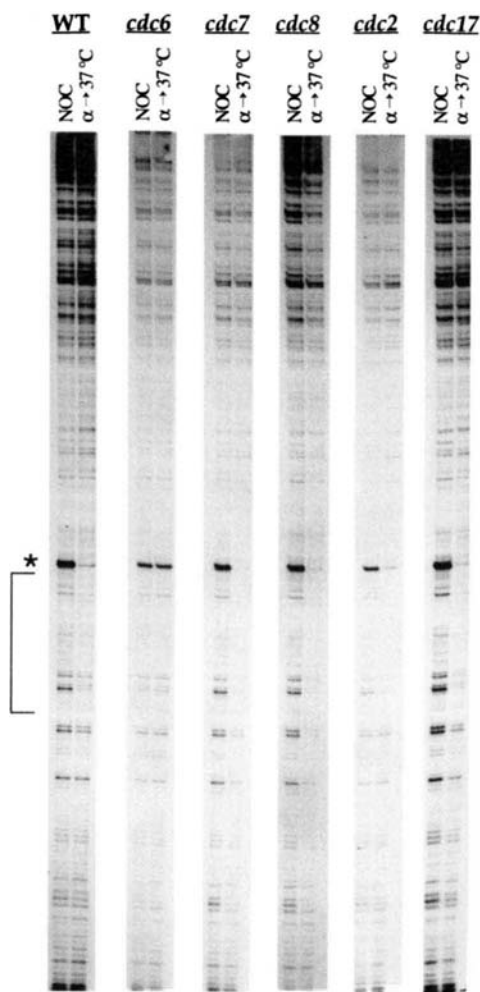


FIG. 4 Pre-RCs are normal in other cell-cycle mutants. Cultures of strains congenic to A364A (see Fig. 3) containing either no *cdc* mutation (wild type) or *cdc6-1*, *cdc7-4*, *cdc8-18*, *cdc2-1* or *cdc17-1* mutants were grown in YEPD medium at the permissive temperature. Half the culture was blocked in G2-M by transferring cells to fresh YEPD medium containing $10 \mu\text{g ml}^{-1}$ nocodazole (NOC) as described above. The other half was blocked in G1 with $10 \mu\text{g ml}^{-1}$ synthetic alpha factor at the permissive temperature until > 90% of the cells were unbudded (~ 2 h) and then raised to 37°C for an additional 2 h ($\alpha \rightarrow 37^\circ\text{C}$). Genomic fingerprints were then performed on each culture. During this experiment, > 90% of the cells remain unbudded after the incubation at the non-permissive temperature.

assembly of pre-RCs. Cells enter G1 with apparently normal kinetics in the absence of Cdc6p suggesting that pre-RCs are not essential for exit from mitosis. These cells, however, fail to replicate their DNA providing evidence that pre-RCs are required for DNA synthesis (S.P. and K.N., unpublished data).

To examine the role of Cdc6p in the maintenance of the pre-replicative state, pre-RCs were first allowed to form in both a wild-type and *cdc6* temperature-sensitive mutant strain (*cdc6^{ts}*) by arresting cells in G1 with alpha factor mating pheromone at the permissive temperature. The cultures were then raised to the restrictive temperature and 2- μm origin chromatin was examined in both strains. Pre-RCs persisted after temperature shift in the wild-type strain (Fig. 3b), but gradually disappeared in the *cdc6^{ts}* strain (Fig. 3c). Thus, pre-RCs are thermolabile in a *cdc6^{ts}* mutant

indicating that Cdc6p is required for both their establishment and maintenance. These data suggest that either pre-RCs are inherently unstable and require constant Cdc6p action for their maintenance or, perhaps more likely, that Cdc6p is an essential component of pre-RCs.

In many other cell-cycle mutants, including several directly involved in DNA replication, replication origins are in the post-replicative state when cells are blocked at the non-permissive temperature¹. This could be because arrest occurs after origin firing and, consequently, after loss of pre-RCs. Alternatively, these gene products could be directly required for pre-RCs *in vivo*. The experiment in Fig. 4 demonstrates that pre-RCs are thermolabile in a *cdc6* mutant but not in *cdc7*, *cdc8*, *cdc2* and *cdc17* mutants. These genes encode, respectively, a protein kinase required for the G1-S transition, thymidylate kinase, the catalytic subunit of DNA polymerase III and the catalytic subunit of DNA polymerase I²⁰. Thus far, *cdc6* mutants are the only mutants we have found in which pre-RCs are thermolabile (J.C., C.S. and J.D., unpublished data).

Loss of Cdc6p in G1 either by promoter shut off (Figs 1 and 2) or thermal inactivation (Figs 3 and 4) yields origins that are very similar to post-replicative origins that we have previously shown can be reconstituted with purified ORC¹ indicating that ORC is present and able to bind origins during G1 and providing evidence that ORC is also a component of pre-RCs. Thus, we propose that pre-replicative origins are bound by both ORC and a G1-specific factor containing Cdc6p. Recently it has been shown that Cdc6p expressed in insect cells is capable of forming stable complexes with ORC in cell extracts⁴, consistent with this hypothesis.

These experiments are consistent with a two-step model for the regulation of initiation of DNA replication. In the first step, Cdc6p-dependent pre-RCs assemble at origins at the end of mitosis. Whether other proteins such as the MCM proteins recently demonstrated to have a role in DNA replication 'licensing' in *Xenopus* egg extracts²¹⁻²³ are involved in pre-RCs remains to be determined. In the second step, pre-RCs are activated, resulting in the firing of origins. This activation step requires the Cdc7 protein kinase, whose regulatory subunit, Dbf4, interacts with initiation complexes *in vivo*²⁴. Activation also requires the Cdc28 protein kinase associated with any one of six B cyclins²⁵. We suggest that the temporal separation of pre-RC assembly from origin firing may play a critical role in blocking re-replication. □

Received 16 June; accepted 26 October 1995.

- Diffley, J. F. X., Cocker, J. H., Dowell, S. J. & Rowley, A. *Cell* **78**, 303-316 (1994).
- Brown, J. A., Holmes, S. G. & Smith, M. M. *Molec. cell. Biol.* **11**, 5301-5311 (1991).
- Bell, S. P. & Stillman, B. *Nature* **357**, 128-134 (1992).
- Diffley, J. F. X. & Cocker, J. H. *Nature* **357**, 169-172 (1992).
- Rowley, A., Cocker, J. H., Harwood, J. & Diffley, J. F. X. *EMBO J.* **14**, 2631-2641 (1995).
- Micklem, G., Rowley, A., Harwood, J., Nasmyth, K. & Diffley, J. F. X. *Nature* **366**, 87-89 (1993).
- Foss, M., McNally, F. J., Laurenson, P. & Rine, J. *Science* **262**, 1838-1844 (1993).
- Bell, S. P., Kobayashi, R. & Stillman, B. *Science* **262**, 1844-1849 (1993).
- Li, J. J. & Herskowitz, I. *Science* **262**, 1870-1874 (1993).
- Loo, S. et al. *Molec. Biol. Cell* **6**, 741-756 (1995).
- Fox, C. A., Loo, S., Dillin, A. & Rine, J. *Genes Dev.* **9**, 911-924 (1995).
- Hartwell, L. H. *J. molec. Biol.* **104**, 803-817 (1976).
- Bueno, A. & Russell, P. *EMBO J.* **11**, 2167-2176 (1992).
- Hogan, E. & Koshland, D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3098-3102 (1992).
- Platt, S., Lengauer, C. & Nasmyth, K. *EMBO J.* **14**, 3788-3799 (1995).
- Liang, C., Weinreich, M. & Stillman, B. *Cell* **81**, 667-676 (1995).

- Kelly, T. J. et al. *Cell* **74**, 371-382 (1993).
- Zwerschke, W., Rottjakob, H.-W. & Kuntzel, H. *J. biol. Chem.* **269**, 23351-23356 (1994).
- Cherest, H., Nguyen, N. T. & Surdin Kenjan, Y. *Gene* **34**, 269-81 (1985).
- Campbell, J. L. & Newlon, C. S. in *The Molecular Biology of the Yeast Saccharomyces; Genome Dynamics, Protein Synthesis and Energetics* (eds Broach, J. R., Pringle, J. R. & Jones, E. W.) 41-146 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Chong, J. P. J., Mahbubani, H. M., Khoo, C.-Y. & Blow, J. J. *Nature* **375**, 418-421 (1995).
- Madine, M. A., Khoo, C.-Y., Mills, A. D. & Laskey, R. A. *Nature* **375**, 421-424 (1995).
- Kubota, Y., Mimura, S., Nishimoto, S.-I., Takisawa, H. & Nojima, H. *Cell* **81**, 601-609 (1995).
- Dowell, S. J., Romanowski, P. & Diffley, J. F. X. *Science* **265**, 1243-1246 (1994).
- Schwob, E., Bohm, T., Mendenhall, M. D. & Nasmyth, K. *Cell* **79**, 233-244 (1994).
- Rose, M. D., Winston, F. & Hieter, P. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, New York, 1989).

ACKNOWLEDGEMENTS. We thank J. Nicholson for photography, I. Goldsmith for oligonucleotide synthesis and G. Evan for peptide synthesis. We also thank T. Hunt for critical reading of this manuscript.

Requirement of mammalian DNA polymerase- β in base-excision repair

Robert W. Sobol*, Julie K. Horton†, Ralf Kühn‡, Hua Gu‡§, Rakesh K. Singhal*, Rajendra Prasad*, Klaus Rajewsky‡ & Samuel H. Wilson*§

* Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555-1068, USA

† Laboratory of Cell Biology, Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555-1068, USA

‡ Institute of Genetics, University of Cologne, Weyertal 121, D-50931 Cologne, Germany

SYNTHESIS of DNA by DNA polymerase- β is distributive on single-stranded DNA templates, but short DNA gaps with a 5' PO₄ in the gap are filled processively to completion^{1,2}. *In vitro* studies have suggested a role of β -polymerase in different types of DNA repair³⁻⁹. However, the significance of these studies to the *in vivo* role of β -polymerase has remained unclear. Because genetic studies are essential for determining the physiological role of a gene, we established embryonic fibroblast cell lines homozygous for a deletion mutation in the gene encoding DNA polymerase- β . Extracts from these cell lines were found to be defective in uracil-initiated base-excision repair. The β -polymerase-deleted cells are normal in viability and growth characteristics, although they exhibit increased sensitivity to monofunctional DNA-alkylating agents, but not to other DNA-damaging agents. Both the deficiency in base-excision repair and hypersensitivity to DNA-alkylating agents are rescued following stable transfection with a wild-type β -polymerase minitransgene. These studies demonstrate that β -polymerase functions specifically in base-excision repair *in vivo*.

Transgenic mice with a heterozygous germline deletion mutation in the 5' end of the DNA polymerase- β (β -pol) gene have been generated¹⁰. When bred, these mice fail to produce offspring that are homozygous for the β -pol deletion mutation, and embryos harbouring the homozygous deletion mutation do not survive beyond day 10.5 of gestation¹⁰. To evaluate cell viability and base-excision repair (BER) capacity in a β -pol null background, primary cultures of fibroblasts were established from embryos of wild-type, heterozygous deletion, and homozygous deletion mice at 10 days gestation. The frequency distribution of the three β -pol genotypes was consistent with Mendelian segregation (Fig. 1a). Primary cultures and the subsequent post-crisis cell lines exhibited a fibroblast-like appearance. No alteration in growth rate was observed in primary cultures from any of the β -pol deletion genotypes, and all established post-crisis cell lines were contact inhibited, regardless of the β -pol genotype. Expression of full-length (*M*_r 39,000 (39 K)) β -pol could be detected by immunoblotting of extracts from wild-type and heterozygous fibroblasts, but not from homozygous β -pol gene deletion fibroblasts (Fig. 1b; lanes 5 and 10). Reduced β -pol expression was observed in the heterozygous cells (Fig. 1b; lanes 4, 6, 9 and 11) as compared to cells from wild-type littermates (Fig. 1b; lanes 3, 7 and 8).

To facilitate analysis of the β -pol deletion genotype with regard to cell growth, DNA-repair capacity and mutagen sensitivity, wild-type (+/+), β -pol-deleted (-/-) and β -pol heterozygous (+/-) transformed cell lines were established by transfection with an expression vector for the SV40 large-T antigen¹¹. XL-PCR analyses were used to examine wild-type and mutant β -pol alleles in the cell lines (Fig. 2a). Genomic DNA from wild-type cells (lane 1), but not from β -pol-deleted cells (lane 2), yielded the 6.5-kb amplified DNA fragment expected from the wild-type gene (Fig. 2b). Amplification

(across the 'deleted' region) of genomic DNA isolated from wild-type and β -pol mutant cells resulted in amplicon sizes of 6.8 and 5.6 kb, respectively (Fig. 2a, c). The internal *Bam*HI site in the β -pol mutant allele was confirmed by *Bam*HI digestion of the amplified product (data not shown) and the loss of the 5.6-kb amplification product, when amplification was carried out using *Bam*HI-predigested genomic DNA (Fig. 2c). These results were confirmed by Southern blot analysis (data not shown). Expression of β -pol mRNA was evaluated by RT-PCR. As anticipated, a 774-base-pair fragment was amplified from oligo(dT)-directed cDNA prepared from total RNA of the wild-type cells; no amplification product was observed in identically prepared samples from β -pol-deleted cells (Fig. 2d). Equal amplification of β -actin mRNA was observed in both cell lines (Fig. 2d; lanes 3 and 4).

The BER capacity of extracts from wild-type (+/+) cells (Fig. 3a; lane 1) was measured by following the repair of a

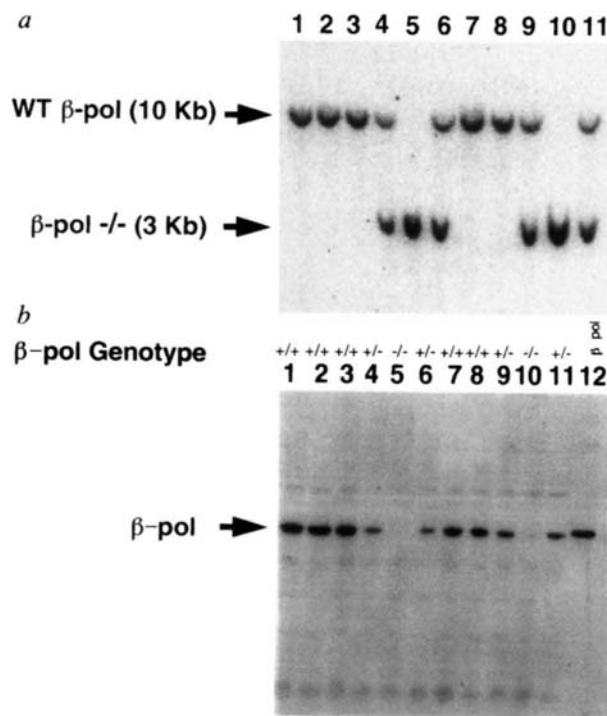


FIG. 1 β -pol genotype identification and protein expression in embryonic fibroblasts. **a**, Southern blot analysis of *Bam*HI-digested genomic DNA isolated from primary fibroblast cultures derived from two control embryos (lanes 1 and 2) or a litter of 9 embryos from β -pol +/- heterozygotes (β T14A). The probe used was described earlier¹⁰. The wild-type β -pol gene exhibits a 10-kb band whereas the deleted β -pol gene exhibits a 3-kb band. Homozygous β -pol wild-type cells (+/+) are shown in lanes 1-3, 7 and 8; heterozygous cells (+/-) are shown in lanes 4, 6, 9 and 11; homozygous β -pol deficient cells (-/-) are shown in lanes 5 and 10. **b**, Immunoblot analysis of whole-cell extracts from murine embryonic fibroblasts. Cells (2×10^5) from primary embryonic fibroblast cultures were lysed in SDS-PAGE loading buffer. Fibroblast cell extract (lanes 1-11) or 100 ng recombinant β -pol (lane 12) were resolved by 12.5% SDS-PAGE and immunoblots were probed for β -pol as described⁷. The genotype with respect to DNA polymerase β as determined by Southern blot analysis (Fig. 1a) is indicated across the top of the gel.

METHODS. Primary cultures of wild-type or DNA β -pol-deficient fibroblasts were isolated by microdissection of embryos from heterozygous females (β -pol +/-) at day 10.5 of gestation. Each embryo was washed several times in sterile PBS, minced and trypsinized at 37°C for 30 min. The remaining cell suspension from each was plated in a single well of a 12-well plate in DMEM containing 10% FBS, pen/strep and Glutamax-I. Cultures were treated with trypsin every 2-4 days and maintained at subconfluent levels by reseeding at 2×10^3 cells per cm², as per the protocol used to generate 3T3 cell lines¹³.

§ Present address: NIH/NIAD, Twinbrook II Facility, 12441 Parklawn Drive, Rockville, Maryland 20852, USA.
|| To whom correspondence should be addressed.

uracil-containing 51-bp oligonucleotide⁷. Repair was inhibited by a neutralizing polyclonal antibody to β -pol (Fig. 3a; lane 2), whereas a slight increase in repair was detected following addition of purified β -pol to nuclear extract from wild-type cells (Fig. 3a; lane 3). Importantly, the nuclear extract from β -pol null ($-/-$) cells failed to exhibit any BER activity (Fig. 3a; lane 4), and the addition of purified β -pol to this nuclear extract restored full repair capacity (Fig. 3a; lane 5). As observed for wild-type cells, BER activity with added β -pol was inhibited by the β -pol neutralizing antibody (Fig. 3a; lane 6). Finally, it must be noted that only the polymerization step in the BER pathway was found to be defective. The uracil-DNA glycosylase, AP-endonuclease and DNA ligase activities were functional in both cell types, as revealed by the restoration of full BER activity following addition of purified β -pol. Interestingly, only β -pol could complement the BER activity, as we found no repair following addition of purified DNA polymerase- α , DNA polymerase- ϵ or DNA polymerase- δ plus PCNA (data not shown).

Growth rate analyses of transformed wild-type and β -pol-deleted cell lines indicated no difference in doubling time (approximately 20 hours) (Fig. 3b). To determine the effect of homozygous β -pol null mutation on the sensitivity of cells to DNA-damaging agents, we examined cell survival following treatment with a variety of DNA-damaging agents. An increase in sensitivity to the monofunctional alkylating agents methyl methanesulphonate (MMS), ethyl methanesulphonate (EMS), *N*-nitroso-*N*-methylurea (MNU), and *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) was observed in the β -pol-deleted cells as compared to wild-type cells (Fig. 3c-f). Similar results were obtained regardless of whether cells were analysed at 5 days or 24 hours after treatment (data not shown). In addition, using the same clonal cell lines, a differential sensitivity to MMS was observed in a clonogenic assay in which control and MMS-treated cells were grown

in soft agar¹¹ (Fig. 3g). Essentially identical results to those in Fig. 3c-f have been observed in four sets of β -pol $+/+$ and $-/-$ cell lines (data not shown). However, no difference in sensitivity to several other DNA-damaging agents was observed, including ultraviolet radiation, γ -irradiation, H_2O_2 , cisplatin, and nitrogen mustard (Fig. 3h,i and data not shown). These latter observations suggest that any BER-repairable damage that is induced by ultraviolet light, γ -rays or H_2O_2 is not lethal. Further, treatment with *O*⁶-benzylguanine, a potent alkyltransferase inhibitor, does not affect the sensitivity differential of MNU (Fig. 3d). As such, it is unlikely that the observed alkylation agent sensitivity differential is the result of differences in alkyltransferase expression (Mer^+ versus Mer^-). Finally, the β -pol $+/+$ and β -pol $-/-$ cell lines are both found to exhibit normal mismatch repair capacity *in vitro* (R. W. S., T. Kunkel and S. H. W., unpublished observation).

The studies described here indicate that β -pol is responsible for the repair of DNA damage conferred by monofunctional DNA-alkylating agents. To verify this conclusion further, we conducted genetic complementation experiments using the expression vector pRS729, encoding a mammalian β -pol minigene. The control vector pCDNA3 (Invitrogen) or pRS729 were transfected into M β 16tsA or M β 19tsA cells and stable, G418 resistant clones were selected. Expression of the recombinant, epitope-tagged β -pol (40 K) was essentially the same as the endogenous protein (Fig. 4a). This resulted in two new cell lines in which (1) β -pol expression as compared to the wild-type M β 16tsA cells (or 16cDNA cells, not shown) was doubled (16/729.B1 cells; Fig. 4a, compare lanes 1 and 3); and (2) β -pol expression in the deleted M β 19tsA cells was returned to near normal levels (19/729.A4 cells; Fig. 4a, compare lanes 2 and 4). Nuclear extract from 16/729.B1 cells was found to have an increased BER capacity as compared to the wild-type M β 16tsA cells (or 16cDNA cells, not shown) (Fig. 4b, compare lanes 1 and 3); and

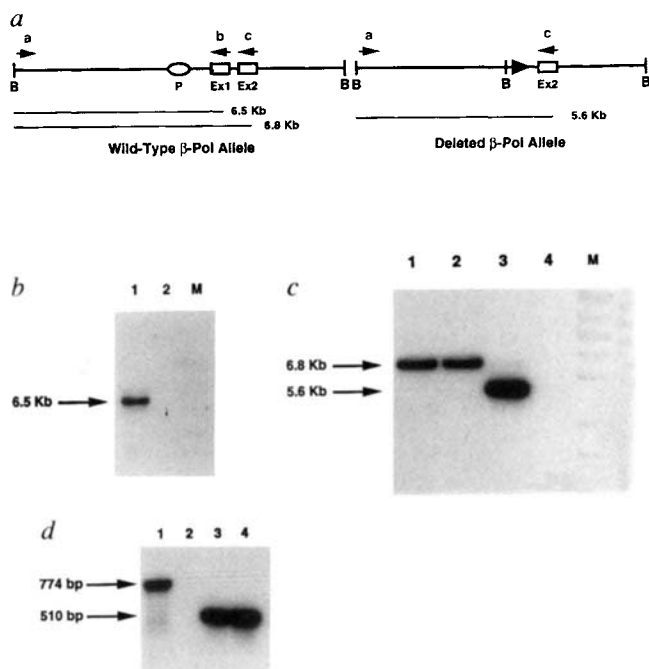
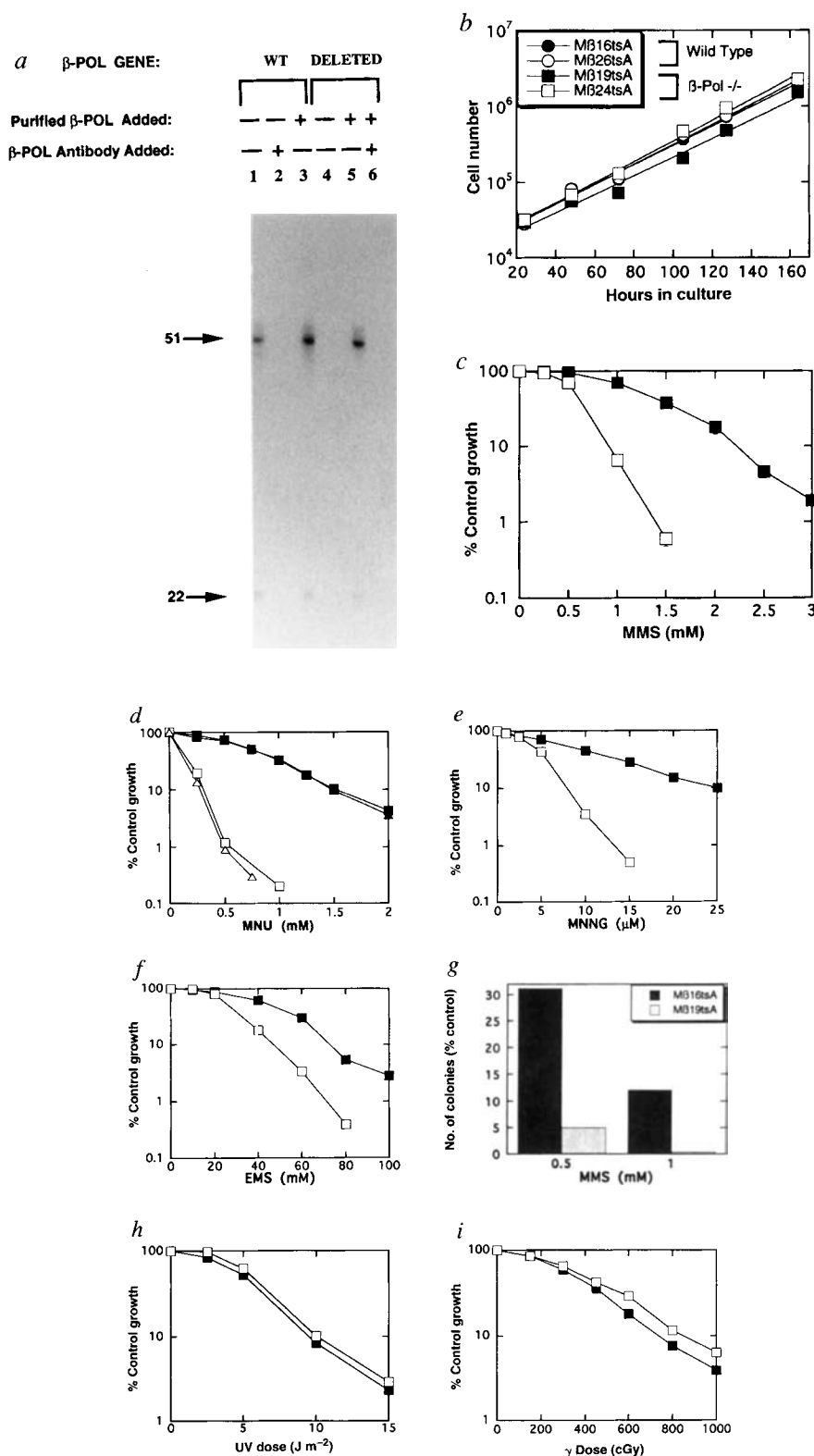


FIG. 2 Genotypic characterization of wild-type and β -pol deficient transformed cell lines. a, Restriction maps for the wild-type (left) and deleted (right) β -pol locus. The open rectangles represent the first and second exons on the β -pol gene, designated Ex1 and Ex2, respectively. The circle represents the β -pol promoter (P); the filled triangle represents a lox P site found in the deleted locus; the primers used for XL-PCR are represented by the arrows labelled a (MBFOR1), b (MBEX1B) and c (MBEX2). BamHI (B) restriction sites are indicated. b, Detection of the wild-type β -pol allele by XL-PCR. The presence of the wild-type allele was determined by XL-PCR amplification of a 6.5-kb fragment using primers MBFOR1 and MBEX1B and genomic DNA from wild-type M β 16tsA cells (lane 1) or β -pol-deficient M β 19tsA cells (lane 2), isolated using a Qiagen tissue kit. c, Detection of the wild-type and deleted β -pol allele by XL-PCR. Utilizing primers MBFOR1 and MBEX2 (a), which span the deleted segment of the β -pol gene, XL-PCR amplification was capable of detecting the wild-type β -pol allele (lanes 1 and 2) as a 6.8-kb fragment and the deleted β -pol allele (lane 3) as a 5.6-kb fragment. Alternatively, XL-PCR amplification was conducted using BamHI-digested genomic DNA from homozygous wild-type or homozygous β -pol-deficient cells (lanes 2 and 4, respectively). d, Analysis of β -pol mRNA by RT-PCR results in the amplification of a 774-bp fragment from wild-type cells (lane 1) but no amplification was observed from the β -pol-deficient cells (lane 2). Identical cDNA preparations were also used to amplify β -actin from wild-type M β 16tsA cells (lane 3) and β -pol-deficient M β 19tsA cells (lane 4). Reactions (b-d; 10–20% total reaction volume) supplemented with [α ³²P]dCTP were separated by agarose gel electrophoresis and transferred to Hybond N⁺ by alkaline transfer¹⁴. Bands were detected by overnight exposure to film at -70°C . Numbers on the left side of the blot indicate the size of the bands. In b and c, lane M contains [^{32}S]DNA markers (Amersham).

METHODS. Primary cultures of wild-type or DNA β -pol-deficient fibroblasts were transfected with the DNA plasmid construct ptsA58H, which contains coding sequences for both a hygromycin-resistance gene and tsA58, a temperature-sensitive SV40 Tag¹¹. Cells were selected in the presence of 80 $\mu\text{g ml}^{-1}$ hygromycin for 21 days, and further selected for 14 days by limiting dilution (0.3 cells per well) in 96-well microtitre plates. Transformed fibroblast cell lines were maintained in DMEM, 10% FBS, pen/strep, Glutamax-I and 80 $\mu\text{g ml}^{-1}$ hygromycin in a 10% CO_2 incubator at 34°C . XL-PCR and RT-PCR reaction conditions available upon request via e-mail at rwsobol@scms.utmb.edu.

FIG. 3 Phenotypic comparison of wild-type and β -pol ($-/-$) fibroblast cell lines. **a**, Uracil-initiated base-excision repair in extracts from fibroblast cell lines. Nuclear extract from wild-type fibroblast cells (lanes 1–3) was reacted alone (lane 1) or in the presence of a neutralizing antibody to β -pol (lane 2) or upon addition of purified, recombinant β -pol (lane 3); whereas nuclear extracts from β -pol-deficient fibroblast cells (lanes 4–6) were reacted alone (lane 4), upon addition of purified, recombinant β -pol (lanes 5 and 6) and in the presence of a neutralizing antibody to β -pol (lane 6) with a 51-base-pair synthetic DNA substrate containing a single G:U base pair as described⁷. Base-excision repair was determined by incorporation of [α -³²P]dCMP exclusively to replace uracil⁷. **b**, Comparison of the growth rate of transformed wild-type and β -pol-deficient fibroblast cell lines. Cells were seeded (40,000 cells per well) in 6-well dishes and grown at 34°C in a 10% CO₂ incubator in growth medium (DMEM, 10% FBS, pen/strep, Glutamax-I), with medium changes daily. For each cell line, cells were counted at 24 h intervals for up to 7 days. Cell numbers were determined by a nuclear lysis protocol¹⁵. Results are expressed as the total number of cells per well, average of three wells. Sensitivity of wild-type and β -pol-deficient fibroblast cell lines to **c**, MMS; **d**, MNU $\pm$ O⁶-benzylguanine (25 μ M)¹⁶; **e**, MNNG; **f**, EMS; **g**, UV and **i**, gamma rays. Cells (wild-type M β 16tsA, clone 1B5; $\blacksquare$; + O⁶-benzylguanine $\blacktriangle$; and β -pol deficient M β 19tsA, clone 2B2; $\square$; + O⁶-benzylguanine $\triangle$) were exposed to the DNA-damaging agents as described below. Results shown in **c** (mean and standard error, $n = 4$) are expressed as the number of MMS-treated cells relative to the cells in control wells (% control growth). **g**, Graphical representation of the number of clones of each cell line showing the ability to grow in soft agar following treatment with MMS as compared to untreated controls.

METHODS. Cytotoxicity was determined by growth inhibition assays. Cells were exposed to the DNA-damaging agents MMS, EMS, MNU, MNNG, cisplatin (not shown), H₂O₂ (not shown) and nitrogen mustard (not shown) at varying doses 24 hours after seeding (40,000 cells per well) in 6-well dishes. Cells (triplicate wells) were treated for 1 hour at 34°C in a 10% CO₂ incubator with serial dilutions of mutagen in antibiotic-free growth medium. For UV irradiation (254 nm – UVC), cell monolayers were washed twice with Hanks balanced salt solution (HBSS) then exposed to UV radiation (fluence 37.5 μ W cm⁻²) for the required length of time. Growth medium was then added to each well. For γ -irradiation, cells in the presence of the growth medium, were exposed to a ¹³⁷Cs source in a Mark I irradiator model 68 (J. L. Shepard and Associates) at a dose rate of 7.33 Gy min⁻¹ for varying lengths of time. After irradiation, the cells were washed once with HBSS and fresh medium was added to each well. For each experiment, cells were counted after 5 days (untreated cells were found to be <80% confluent). Cell numbers were determined by a nuclear lysis protocol¹⁵. Results are the average of (or are representative of) at least two separate experiments using at least two separate cell lines for each β -pol genotype, and were determined as the number of treated cells relative to the cells in control wells (% control growth). For clonal analysis, cells were trypsinized and grown in soft agar as described¹¹. Cell clones were counted 2 weeks later and compared to untreated cells grown under identical conditions.



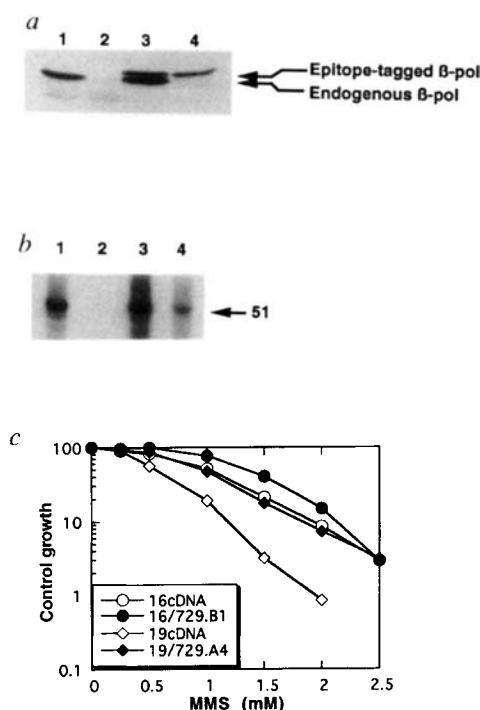


FIG. 4 Genetic complementation of BER deficiency with expression of β -pol minigene. Wild-type and β -pol-deficient cell lines were transfected with pCDNA3 (Invitrogen) or the β -pol expression vector pRS729, respectively. Cells were selected in the presence of $80 \mu\text{g ml}^{-1}$ hygromycin and $600 \mu\text{g ml}^{-1}$ G418 for 28 days, and further selected for 14 days by limiting dilution (0.3 cells per well) in 96-well microtitre plates. Clones harbouring pRS729 were further screened for expression of β -pol by immunoblot analysis. **a**, Immunoblot analysis of nuclear extract with β -pol-specific polyclonal antibody. $250 \mu\text{g}$ of nuclear extract from wild-type M β 16tsA cells (lane 1), β -pol deficient M β 19tsA cells (lane 2), 16/729.B1 cells (lane 3) or 19/729.A4 cells (lane 4) was resolved by 12.5% SDS-PAGE and probed for β -pol as described⁷. A photograph of the autoradiogram is shown. Epitope-tagged and endogenous β -pol are indicated by the arrows. **b**, Uracil-initiated BER *in vitro*. Nuclear extract from wild-type M β 16tsA cells (lane 1), β -pol-deficient M β 19tsA cells (lane 2), 16/729.B1 cells (lane 3) or 19/729.A4 cells (lane 4) were reacted with a 51-base-pair synthetic DNA substrate as described in Fig. 3a legend. **c**, Alkylating agent sensitivity. Cells were exposed to the DNA damaging agent MMS as described in the legend to Fig. 3b. 16cDNA and 19cDNA cells are M β 16tsA and M β 19tsA cells, respectively, stably transfected with the control vector pCDNA3; 16/729.B1 and 19/729.A4 cells are M β 16tsA and M β 19tsA cells, respectively, stably transfected with the β -pol expression vector pRS729. The β -pol expression vector pRS729 will be described elsewhere (R.W.S. and S.H.W. manuscript in preparation).

nuclear extract from 19/729.A4 cells was now found to be proficient in BER (Fig. 4b, compare lanes 2 and 4). This genetic complementation is also observed *in vivo*, as the hypersensitivity of the β -pol-deleted cells to MMS was completely reversed in the 19/729.A4 cells (Fig. 4c). Further, 16/729.B1 cells exhibit an increased resistance to MMS, suggesting that β -pol may be the rate-limiting factor in the BER pathway *in vivo*.

It has been previously thought that β -pol functions in repair synthesis in a manner analogous to the *Escherichia coli* PolA

encoded DNA polymerase I, which also is required for strand-displacement synthesis following the removal of RNA primers from newly replicated DNA. PolA mutants are sensitive to a wide variety of DNA-damaging agents including MMS and ultraviolet radiation¹². Our studies with β -pol null mutants, however, provide no support for a role of β -pol in either DNA replication or the repair of damage from ultraviolet light (for example, nucleotide-excision repair), as β -pol mutant cells grow normally and are not sensitive to ultraviolet radiation. Rather surprisingly, we find that mammalian β -pol has a highly specific role in BER *in vivo*. Whereas DNA polymerases- α , δ and ϵ are unable to substitute for β -pol in the BER reaction *in vitro*, similar genetic studies will be required to delineate their role precisely in DNA repair *in vivo*. The apportionment of β -pol to BER presumably reflects the importance of this repair pathway in maintaining the fidelity and integrity of genomic DNA in mammalian cells. □

Received 2 August; accepted 31 October 1995.

1. Singhal, R. K. & Wilson, S. H. *J. biol. Chem.* **268**, 15906–15911 (1993).
2. Prasad, R., Beard, W. A. & Wilson, S. H. *J. biol. Chem.* **269**, 18096–18101 (1994).
3. Wiebauer, K. & Jiricny, J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5842–5845 (1990).
4. Matsumoto, Y. & Bogenhagen, D. F. *Molec. cell. Biol.* **11**, 4441–4447 (1991).
5. Dianov, G., Price, A. & Lindahl, T. *Molec. cell. Biol.* **12**, 1605–1612 (1992).
6. Wilson, S. H. in *The Eukaryotic Nucleus: Molecular Structure and Macromolecular Assembly* Vol. 1 (eds Strauss, P. R. & Wilson, S. H.) 199–234 (Telford, Caldwell, N.J., 1991).
7. Singhal, R. K., Prasad, R. & Wilson, S. H. *J. biol. Chem.* **270**, 949–957 (1995).
8. Fornace, A. J. Jr, Zmudzka, B., Hollander, M. C. & Wilson, S. H. *Molec. cell. Biol.* **9**, 851–853 (1989).
9. Srivastava, D. K. et al. *J. biol. Chem.* **270**, 1–7 (1995).
10. Gu, H. et al. *Science* **265**, 103–106 (1994).
11. Sell, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 11217–11221 (1993).
12. Kornberg, A. & Baker, T. A. *DNA Replication*, 2nd edn, (Freeman, New York, 1991).
13. Todaro, G. & Green, H. *J. Cell Biol.* **17**, 299 (1963).
14. Kriegler, M. *Gene Transfer and Expression*, 211–213 (Freeman, New York, 1990); *Current Protocols in Molecular Biology* (Wiley, New York, 1995).
15. Butler, W. B. *Analyt. Biochem.* **141**, 70–73 (1984).
16. Dolan, M. E., Moschel, R. C. & Pegg, A. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5368–5372 (1990).

ACKNOWLEDGEMENTS. We thank W. A. Beard, R. S. Lloyd, S. Mitra, S. Prakash and B. Van Houten for critical reading of the manuscript, R. Baserga for pA58H, S. Widen for pMBAM18 (a plasmid containing β -pol genomic DNA used to determine DNA sequence information for the design of the PCR primers), C. Sell for helpful discussions and M. Yakes and Y. Chen for help with XL-PCR method development and analysis. Supported in part by grants from the NIH and Robert A. Welch Foundation (to S.H.W.). The work in Cologne was supported by the Deutsche Forschungsgemeinschaft.

ERRATUM

A nesting dinosaur

Mark A. Norell*, James M. Clark†, Luis M. Chiappe‡ & Demberelyin Dashzeveg§

* Department of Vertebrate Paleontology, and
 ‡ Department of Ornithology, American Museum of Natural History,
 79th Street at Central Park West, New York, New York 10024-4192, USA
 † Department of Biological Sciences, George Washington University,
 Lisner Hall 307, 2023 G Street NW, Washington DC 20052, USA
 § Geological Institute, Mongolian Academy of Sciences, Ulaan Baatar 11,
 Mongolia

Nature **378**, 774–776 (1995)

A PRODUCTION problem after the galley proof stage caused confusion in the authors' affiliations for this Letter in the 21/28 December issue. The affiliations shown above are correct, with M.A.N. at the American Museum of Natural History.

Duplicate DNA

Ample product listings for improved DNA amplification tools — including an automated PCR system, primer pair analysis software, an automated PCR-DNA fragment analysis system and various consumable amplification products.

AMERSHAM has announced the release of the **Vistra Fluorescence 5'-oligolabelling kit** for fast, fluorescent 5'-oligolabelling (*Reader Service No. 100*). The kit is said to be suitable for PCR primer labelling and gel-shift assays. It is designed to attach a single fluorescein dye molecule to the 5' end of any oligonucleotide quickly and easily. For use with the FluorImager SI, this oligonucleotide-based assay requires no film exposures or darkroom processing and provides rapid, quantitative results.

Epicentre Technologies now offers MMLV-RT (Moloney murine leukaemia virus reverse transcriptase), for **first-strand cDNA synthesis** in the construction of cDNA libraries, for production of templates for RT-PCR amplifications, as well as for 5'-end transcript mapping via primer extension analysis (*Reader Service No. 101*). The manufacturer states that under standard reaction conditions, as little as five units of the company's MMLV-RT per microgram of RNA template synthesizes as much full-length cDNA as is obtained with much larger unit quantities (up to 300 units) of other MMLV-RTs, including modified (RNase H minus) enzymes. MMLV-RT is supplied with a 10× reaction buffer and is certified to be RNase-free.

Tepnel's Daras **automated PCR system** is designed to allow the manipulation and detection of PCR products to be automated for highly reproducible results (*Reader Service No. 102*). The system combines the amplification potential of PCR with the sequence-specific capture, manipulation and detection of DNA on a solid phase, bringing speed, automation and sensitivity to DNA analysis and manipulation. This new technology has the potential to be used in combination with PCR in a wide variety of molecular biology research applications — whenever there is a need to identify nucleic acids accurately, in a wide range of samples, including bacteria, tissue and cell cultures.

Roche Diagnostic Systems has announced the release of Cobas Amplicor, a fully **automated DNA/RNA amplification and detection system**, which is designed to support up to 24 different amplification reactions and 24 detection procedures simultaneously (*Reader Service No. 103*). The system is said to offer clinical laboratories a range of PCR-based assays for the detection of important infectious diseases,



Roche Diagnostic Systems' Cobas Amplicor system.

including the hepatitis C virus (HCV), *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, and *Mycobacterium tuberculosis*. This single-unit, benchtop system amplifies up to 24 separate samples, while simultaneously performing detections on 24 previously amplified samples. Automated analysis supports true multiplexing — the ability to perform multiple amplifications and/or detections on a single sample. Once detection begins, they are completed at a rate of 50 per hour. Through a feature called 'reflex testing', which is included with the instrument, the operator can program the system to perform automatically an additional round of tests on a sample, based on the results of the initial test.

Software

Cambio has announced the arrival of Oligo 5.0, the latest version of the company's software package for **analysing primer pairs** (*Reader Service No. 104*). It provides researchers using PCR and DNA sequencing with a tool for selecting optimal PCR primer pairs from any DNA/RNA sequence. The search for PCR primers is said to be fast and easy. All search-selected primer pairs are listed, and selected primer pairs are com-

pared in a table along with position number, optimal annealing temperature, PCR products size and the G+C content of the PCR product. Users can sort by any column and click on any pair to show the DNA amplification window. One of the new features is 'nested primer selection', which enables the user to find all cross-compatible interest generated in a search. Another is 'primer T_m matching', which automatically adjusts the length of compatible PCR primer pairs that have disparate T_ms. When this feature is active, nucleotides are dropped from the 3' end of the higher T_m oligonucleotide to bring T_ms into agreement. Screening out the stable (and very unstable) 3' ends is possible with the 'stability window'. To complement the moderate or low 3' stability necessary for specificity, a 5' G+C clamp

filter ensures that all selected oligos have the desired, highly stable 5' end. Another useful feature enables users to download primer selections and to send them immediately by fax or e-mail to a synthesis supplier.

Completely **automated PCR amplification** is now offered with Catalyst PCR software from Perkin-Elmer (*Reader Service No. 105*). Designed to run on the company's ABI 800 Catalyst Molecular Biology LabStation, the software provides a flexible system for integrated pipetting and PCR amplification. The software controls all pipetting, thermal cycling and pooling to be performed in one operation. Engineered for versatility, the software enables users to perform up to 96 PCR reactions with up to 16 primer sets per run. When equipped with this software, the LabStation is said to perform PCR reactions in as little as 5 µl. It is designed to be easy to use and to ensure fast completion of all genotyping, RAPD and SSCP applications. The software is said to be equally efficient for amplifying multiple templates with one primer set, or several templates with multiple primer sets. Another feature, post-PCR sample pooling, allows separate amplification of a DNA template with multiple primer

PRODUCT REVIEW

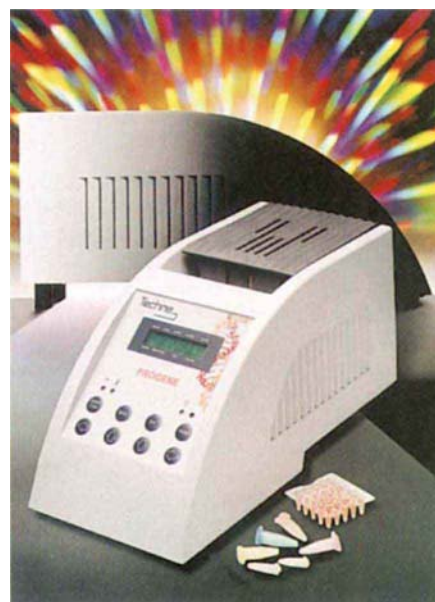
sets followed by automatic pooling in preparation for analysis with ABI 672 Genescan software. Because the requirements of each run can vary, Catalyst PCR software allows users to customize variables, such as PCR temperature, reagent quantity and the duration of each step in a cycle. Once the parameters are entered, the LabStation takes over completely, performing up to six PCR runs back-to-back, with no supervision required.

Temperature control

Stuart Scientific has responded to the evolving needs of the research community by improving the design of its SPCR1 Gene-Tech Cryo **thermal reactor** (Reader Service No. 106). It has a more compact and simplified housing. Moreover, the software has been improved to prompt the operator during programming, and to give a choice of two temperature cut-off points, within 1.0 °C or 1.5 °C of a desired temperature. During operation, when the temperature is increasing or decreasing towards a set temperature, there is a cut-off point at which the temperature change starts to slow before plateauing. This is designed to minimize temperature overshoot or undershoot. The manufacturer states that by using the optional 1.5 °C cut-off point, the overall speed of a 30-cycle regime can be increased by 50 per cent. There is also built-in Peltier refrigeration

that covers the temperature range from 4 °C to 99 °C. It has 48 linkable programmes with 28 segments each, of which 40 have automatic control so that only dwell times have to be specified (all other times are automatically set to minimum).

The new Progene **thermal cycler** from Techne is designed to have numerous applications, in molecular biology and diagnostic laboratories and is suitable for use with diagnostic kits (Reader Service No. 107). Its features include 'user friendly' software that can accommodate all the latest PCR protocols, such as 'Touchdown', 'RAPD' and 'Long PCR'. According to the manufacturer, a typical 25-cycle programme can be completed in about an hour, with a cooling/heating rate of 2 °C/3 °C per second. Two models are available for 20 × 0.5-ml microcentrifuge tubes and 25 × 0.2-ml micro-tube/Costar 5 × 5 Thermowell plates. Each unit is fitted with a heated lid as standard and comes complete with a two-year warranty. Progene operates over the temperature range 4 °C to 99 °C and uses Peltier technology, giving it a stated block uniformity of better than ±0.5 °C for both internal and (through the external probe) external control and calibration. Other features include interchangeable blocks, automatic restart, incremental/decremental timing, incremental/



The Progene thermal cycler from Techne.

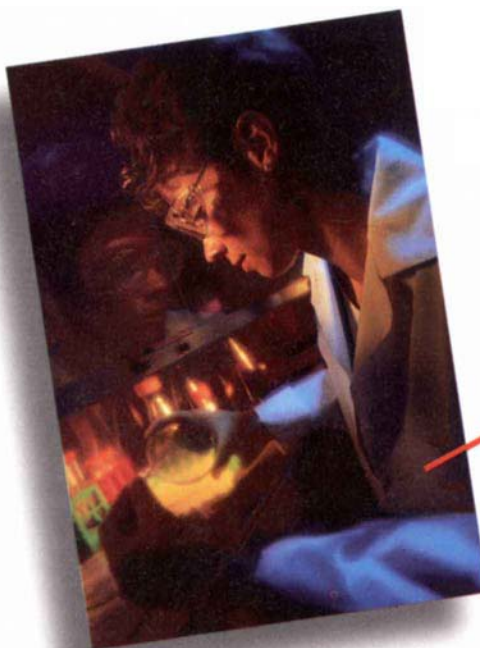
decremental temperature programming, an RS232 computer linking and a manual pause button.

Separate and quantitate

Amicon offers a new, 24-page publication for researchers that discusses creative solu-

Doris Dixon

- Presently researching protein expression and DNA/RNA isolation and purification systems.
- Published in *Gene*, *Journal of Immunology*, and *Molecular Biology Cell*.
- Presented abstracts at Symposium for Protein Society and ASBMB/DBC-ACS Joint Meeting.
- Worked four years as Senior Research Scientist in the Department of Molecular and Cellular Biology at a major pharmaceutical company.
- Worked ten years at various academic research laboratories including University of Missouri, University of Texas and Wadsworth Research Institute at the New York State Department of Health.



"You demand only the finest molecular biology products. I demand the same thing."

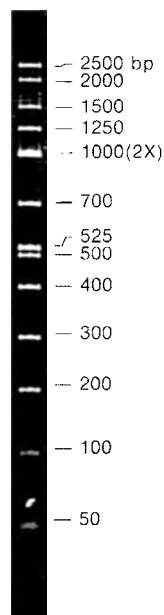
READER ENQUIRY NO. 477

tions to bioseparation problems (*Reader Service No. 108*). It describes 19 procedures that are designed to save time and effort and features the company's range of membrane ultrafiltration devices. Featured applications include DNA concentration, desalting and buffer exchange, DNA recovery from agarose gels, protein recovery from dilute solutions, enzyme removal from DNA and purification of PCR products. A short summary of step-by-step procedures and an illustration of the company's centrifugal membrane ultrafiltration devices are included. Referenced materials for procedures are available through the company's 'fax solutions' system, a free, 24-hour-a-day, 7-days-a-week, document retrieval system.

Advanced Molecular Systems has announced the availability of its FingerprinterCE system for, **automated PCR-DNA fragment analysis** in 15 minutes (*Reader Service No. 109*). The completely non-isotopic method can detect and quantitate as little as one nanogram per millilitre starting concentration of PCR-generated DNA fragments. The system is able to perform multiplex PCR analysis using a single fluorescently labelled molecular probe without hybridization or using multiple fluorescently labelled primers or probes. It is a fully automated, high-speed electrophoresis system uti-

lizing the capillary format and real-time detection. The FingerprinterCE system is equipped with separate photodiode array fluorescence and ultraviolet detectors, and samples directly from 96-well microplates or microcentrifuge tubes.

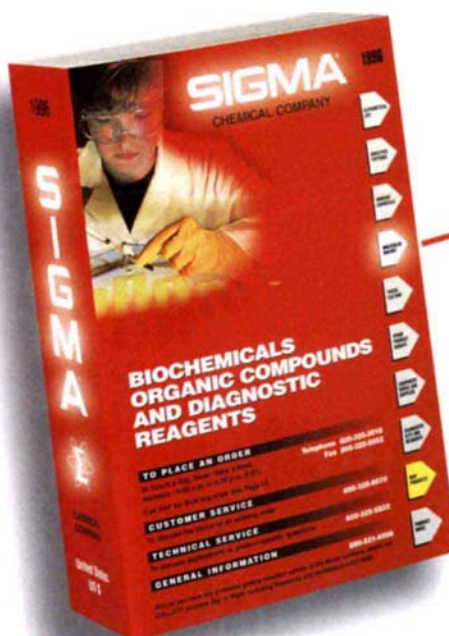
BioMarker EXT Plus is a new **electrophoresis** marker from J. T. Baker, which consists of exact-sized DNA for unambiguous bands (*Reader Service No. XXX*). There is no extraneous DNA or co-migration as found with some markers. With easy-to-identify bands, the marker features a 500/525-bp doublet and a 1000-bp, 2X-intensity band as landmarks. The manufacturer states that this marker is suitable for PCR with a range of 50 to 2,500 bp.



BioMarker EXT.

Consumables

To eliminate aerosols, Tri-Continent has developed a new **range of PCR pipettes and tips** combining an accurate, positive displacement design, a patented tip locking mechanism and longer, narrower tips (*Reader Service No. 110*). The tip plunger creates a seal inside the microsyringe. This prevents any molecules, even in aerosol form, from passing into the barrel of the pipettes and causing cross-contamination. Precision moulding techniques have been used to produce tips that are longer and narrower, enabling operators to pierce oil layers with more accuracy. When the technician has completed the dispensing, the tip and plunger are easily ejected with the release of the operator's thumb from the fully depressed plunger. Multistep procedures can be executed quickly by rotating the control knob at the top of the pipette to select the volume desired. A digital readout displays the volume, and is locked into place as soon as it is set. Volume ranges from 0.5 to 25 μ l are adjustable in 0.1- μ l increments and in 1- μ l increments for volume ranges of 25 to 250 μ l. The disposable Tri-Continent tips and plungers are available pre-assembled as a unit and are biocompatible. Tips can be sterilized before use by autoclaving at 121 $^{\circ}$ C for 15 minutes. The company is now offering positive-



"The best place to find these products is right here."

Like all our molecular biologists, Doris actively participates in ensuring that the products that make it into the Sigma Catalog are the finest available. Everything from Agarose to *in vitro* transcription systems must pass our quality control tests. And just like the Sigma reagents you know and trust, all Sigma Molecular products are unconditionally guaranteed.

For highly-specific, highly-sensitive molecular products and technical support, **Go to the Book.**

Call 1-800-325-3010 (USA/Canada) or 314-771-5750 (outside USA and Canada) for your personal copy.
NATURE • VOL 379 • 11 JANUARY 1996

SIGMA
BIO SCIENCES

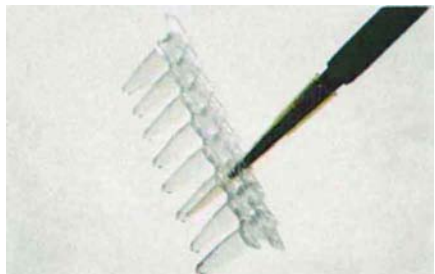


Tri-Continent positive-displacement tips.

displacement tips that are certified to be sterile and free from DNase, RNase, DNA and endotoxin.

Robbins Scientific has introduced a new **single, thin-walled, 0.2-ml PCR tube** to join the company's line of disposable PCR products (*Reader Service No. 111*). Designed for completely oil-free operation, the tube has an attached 'domed' cap that interfaces with the various types of heated lids used by thermal cycler manufacturers to maintain isothermal conditions within the tube. The specially designed cap locks securely after reagent addition to prevent leakage. The thin-walled, 0.2-ml PCR tubes come with a frosted writing surface on both the cap and around the side of the tube to simplify sample identification, and are available non-sterile or radiation-sterilized. Each lot of tubes is tested for RNase and DNase contamination by an independent laboratory and is certified free of these nucleases.

Advanced Biotechnologies has announced the availability of Thermo-Pierce, a new **eight-cap strip** that allows the withdrawal of reagent mix from the tube without removing the cap (*Reader Service No. 112*). The cap strip is designed for sealing the company's range of thin-walled, 0.2-ml tubes, which are used in thermal cycling



Thermo-Pierce's new eight-cap strip.

procedures. Thermo-Pierce should be of interest to laboratories where contamination is an issue. It is said to provide 100 per cent sealing, but allows for the withdrawal of the reaction mix without generating contaminating aerosols. These strips are available ster-

ile or non-sterile and are suitable for use with the 0.2-ml Microstrips, 0.3-ml Omnistrips, Micro-Tube plates, Omni-Tube plates and Thermo-Fast.

Perform RT-PCR quickly, easily and in one tube by using Molecular Bio-Products' HotStart **storage and reaction tubes** (*Reader Service No. 113*). The pre-positioned wax bead enables the researcher to separate reverse transcription (RT) and PCR reactions within the same tube until after the cDNA has been made and is ready for amplification. Simply add the reverse transcription reagents to the tube and during the denaturing step (typically 60 °C) the wax bead melts to cover the RT reaction. While the RT reaction is occurring, the wax overlay will harden. Now add your PCR reagents and when the reaction begins, the wax barrier will melt and rise to the top, allowing the RT and PCR reactions to combine. By eliminating the need to transfer the samples from one tube to another, these tubes significantly reduce the risk of contamination and sample loss. Reaction time can be reduced also, because the transition from reverse transcription to PCR is much faster.



RT-PCR Hotstart.

Literature

Sigma-Aldrich Techware now offers a comprehensive **laboratory manual** that details PCR requirements and techniques (*Reader Service No. 114*). *PCR Primer: A Laboratory Manual* explores the basic requirements of performing PCR and other amplification techniques, as well as advanced PCR applications. It covers issues such as sample preparation, primer design, efficiency, detection of products and quantification. Protocols for a wide range of PCR and amplification techniques are presented

for cloning, sequencing, mutagenesis, library construction and screening, exon trapping, differential display and expression. Protocols include RT-PCR, RNA PCR, LCR, multiplex PCR, panhandle PCR, capture PCR, expression PCR, 3' and 5' RACE, *in situ* PCR, and ligation-mediated PCR. Each protocol is further supplemented with analysis, troubleshooting, and reference information. □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

DNA SEQUENCING

£45 per reaction

- PCR products
- M13 inserts
- plasmid inserts
- single strand DNA
- double strand DNA
- no result? no charge!

By research labs for research labs

Molecular Biology Services
King's College School of Medicine and Dentistry
 Rayne Institute, 123 Coldharbour Lane, London SE5 9NU
 Tel: 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO. 11

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT

PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray
 Phone: +44 191 222 8612
 Fax: +44 191 222 8129
 Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 12

Severn Biotech Ltd

Severn Biotech Ltd, Unit 2, Park Lane,
 Stourport Road, Kidderminster, Worcs DY11 6TJ, U.K.

Custom DNA & RNA Oligo's
 Probes, Primers, Antisense Oligo's, Ribonucleotides, 5' and 3' labelling with Biotin Fluorescein, Rhodamine, Modified bases purification etc.

Complete Peptide Service
 Small Scale Multiple Peptides 5-10mgs or larger-scale 1000mgs using L and D Amino Acids, screening grade of HPLC purified, Phosphorylated, Acetylated Amidated and Protein Carrier Conjugated (KLH, PPD, etc.)

Customer Services
 Comprehensive Custom Service for DNA, RNA and Peptides from SBL

- DNA Synthesis
- RNA Synthesis
- Multi Peptide Synthesis (5-10mgs)
- Larger Scale Peptide Synthesis 1000mgs
- HPLC Purification
- Custom Conjugation to Carriers
- Peptide Sequencing / Mass Spec.

For a Fast Efficient Service Reply Today
Telephone: (01562) 825286 Fax: (01562) 825284

READER ENQUIRY NO. 80